



US009119815B2

(12) **United States Patent**
Glueck et al.

(10) **Patent No.:** US 9,119,815 B2
(45) **Date of Patent:** Sep. 1, 2015

(54) **COMBINED MEASLES-MALARIA VACCINE**

(75) Inventors: **Reinhard Glueck**, Ahmedabad (IN);
Agata Fazio, Catania (IT); **Viviana Giannino**, Catania (IT); **Martin A Billeter**, Zurich (CH)

(73) Assignee: **CADILA HEALTHCARE LIMITED**, Ahmedabad (IN)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **13/318,701**

(22) PCT Filed: **May 3, 2010**

(86) PCT No.: **PCT/IN2010/000287**

§ 371 (c)(1),
(2), (4) Date: **Jan. 23, 2012**

(87) PCT Pub. No.: **WO2010/128524**

PCT Pub. Date: **Nov. 11, 2010**

(65) **Prior Publication Data**

US 2012/0121538 A1 May 17, 2012

(30) **Foreign Application Priority Data**

May 5, 2009 (IN) 1181/MUM/2009

(51) **Int. Cl.**

C12N 15/00 (2006.01)
A61K 39/165 (2006.01)
A61K 39/015 (2006.01)
A61K 38/20 (2006.01)
C12N 7/00 (2006.01)
A61K 35/13 (2015.01)
A61K 39/00 (2006.01)

(52) **U.S. Cl.**

CPC **A61K 39/165** (2013.01); **A61K 38/2013** (2013.01); **A61K 39/0015** (2013.01); **A61K 39/015** (2013.01); **C12N 7/00** (2013.01); **A61K 35/13** (2013.01); **A61K 2039/5256** (2013.01); **A61K 2039/53** (2013.01); **A61K 2039/54** (2013.01); **A61K 2039/55566** (2013.01); **A61K 2039/55588** (2013.01); **A61K 2039/575** (2013.01); **A61K 2039/70** (2013.01); **C12N 2760/18421** (2013.01); **C12N 2760/18434** (2013.01); **C12N 2760/18442** (2013.01); **C12N 2760/18443** (2013.01)

(58) **Field of Classification Search**

CPC A61K 39/165
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,112,749 A 5/1992 Brey, III et al.
5,756,101 A 5/1998 Paoletti et al.
6,214,353 B1 4/2001 Paoletti et al.

2005/0208078 A1 * 9/2005 Hoffman et al. 424/272.1
2005/0265974 A1 12/2005 Pau et al.
2005/0266017 A1 * 12/2005 Druilhe et al. 424/191.1
2006/0127413 A1 6/2006 Sutter et al.
2007/0071726 A1 3/2007 Pau et al.
2007/0088156 A1 4/2007 Pau et al.

FOREIGN PATENT DOCUMENTS

AU	A-16681/97	12/1997
CA	2 507 915 A1	7/2004
EP	0 191 748 A1	8/1986
JP	2010-99041 A	5/2010
WO	94/28930 A1	12/1994
WO	97/06270 A1	2/1997
WO	2009/021931 A1	2/2009
WO	2010/079505 A2	7/2010

OTHER PUBLICATIONS

Li et al (Vaccine vol. 25, pp. 2567-2574, 2007).*
Ballou, W.R., and Cahill, CP (2007). Two Decades of Commitment to Malaria Vaccine Development: GlaxoSmithKline Biologicals. Am. J. Trop. Med. Hyg., 77(6_Suppl), 289-295.
Blatckman, M.J., Whittle H., and Holder AA (1991), Processing of the *Plasmodium falciparum* major merozoite surface protein-1: identification of a 33-kilodalton secondary processing product which is shed prior to erythrocyte invasion. Mol. Biochem. Parasitol., 49(1), 35-44.
Calain, P., and Roux, L. (1988), Generation of measles virus defective interfering particles and their presence in a preparation of attenuated live-virus vaccine. J. Virol., 62 (8):2859-2866.
Cortes, A., Mellomho, M., Masciantonio, R., Murphy, V.J., Reederm J.C., and Anders, R.F. (2005), Allele specificity of naturally acquired antibody responses against *Plasmodium falciparum* apical membrane antigen 1, Infect. Immun., 73: 422-430.
Dilraj, A., Cutts, F.T., de Castro, J.F., Wheeler, J.O., Brown, D., Roth, C., Coovadia, H. M., Bennett, J.V. (2000), Response to different measles vaccine strains given by aerosol and subcutaneous routes to schoolchildren: a randomised trial. Lancet, 355(9206): 798-803.
Enders, I. F., and Peebles, T.C. (1954), Propagation in tissue cultures of cytopathogenic agents from patients with measles. Proc. Soc. Exp. ... Biol. Me'd., 86, 277-286.
Garcia, J.E., Puentes, A., and Patarrroyo, M.E. (2006), Development biology of sporozoite-host interactions in *Plasmodium falciparum* malaria: implications for vaccine design. Clin. Microbiol. Rev., 19(4): 686-707.
Girard, M. P., Reed, Z.H., Friede, M., and Kieny, M.P. (2007), A review of human vaccine research and development: Malaria. Vaccine, 25: 1567-1580.
Griffin, D. (2007) Measles virus. In: Fields Virology, fifth edition, eds.-in-chief Knipe, D.M. & Howley, P. M. Lippincott Williams & Wilkins, Philadelphia PA 19106, USA, 1-53.

(Continued)

Primary Examiner — Albert Navarro

(74) Attorney, Agent, or Firm — Ladas & Parry LLP

(57) **ABSTRACT**

A combined measles-malaria vaccine containing different attenuated recombinant measles-malaria vectors comprising a heterologous nucleic acid encoding several *Plasmodium falciparum* antigens is described. Preferably, it relates to viral vectors that comprise nucleic acids encoding the circumsporozoite (CS) protein of *P. falciparum*, the merozoite surface protein 1 (MSP-1) of *P. falciparum*, and its derivatives (p-42; p-83-30-38) in its glycosylated and secreted forms, and apical membrane antigen1 (AMA1) of *P. falciparum*, in its anchored or secreted form. The viral vector stems from an attenuated measles virus, based on a strain that is used as a vaccine and is efficient in delivering the gene of interest and that binds to and infects the relevant immune cells efficiently.

(56)

References Cited**OTHER PUBLICATIONS**

- Hilleman, M.R. (2002), Current overview of the pathogenesis and prophylaxis of measles with focus on practical implications. *Vaccine*, 20: 651-665.
- Holder AA and Freeman, R.R. (1984), The three major antigens on the surface of *Plasmodium falciparum* merozoites are derived from a single high molecular weight precursor, *J. Exp. Med.*, 160(2): 624-629.
- Martin, A., Staeheli, P. and Schneider, U. (2006), RNA polymerase II-controlled expression of antigenomic RNA enhances the rescue efficacies of two different members of the Mononegavirales independently of the site of viral genome replication. *J. Virol.*, 80, 5708-5715.
- Ovsyannikova I.O., Reid, K.C., Jacobson, R.M., Oberg, A.L., Klee, O.O., Poland, G.A. (2003). Cytokine production patterns and antibody response to measles vaccine. *Vaccine*, 21(25-26), 3946-3953.
- Parks, C. L., Lerch, R. A., Walpita, P., Wang, H., P., Sidhu, M. S., and Udem, S. A (2001), Analysis of the noncoding regions of measles virus strains in the Edmonston vaccine lineage. *J. Virol.*, 75, 921-933.
- Parks, C. L., Lerch, R. A., Walpita, P., Wang, H. P., Sidhu, M. S., and Udem, S. A (2001). Comparison of predicted amino acid sequences of measles virus strains in the Edmonston vanne lineage. *J. Virol.*, 75, 910-920.
- Polley, S.D., Mwangi, T., Kocken, C. H., Thomas, A. W., Dutta, S., Lanar, D. E., Remarque, E., Ross, A., Williams, T.N., Mwambingu, G., Lowe, B., Conway, D.J., and Marsh, K. (2004). Human antibodies to recombinant protein constructs of *Plasmodium falciparum* apical membrane antigen 1 (AMA1) and their association with protection from malaria, *Vaccine*, 23: 718-728.
- Radecke, F., P.- Spielhofer, H. Schneider, K. Kaelin, M. Huber, K. D[delta]tsch, O. Christiansen, and M. Billeter. (1995).Rescue of measles viruses from cloned DNA, *EMBO Journal*, 14: 5773-5784.
- Radecke, F., and M. Billeter, (1997),Reverse genetics meets the nonsegmented negative-strand RNA viruses. *Rev. Med. Virol.*, 7: 49-63.
- Remarque, E.I., Faber, B.W., Kocken, CH. M., and Thomas, A.W. (2008). A diversity-covering approach to immunisation with *Plasmodium falciparum* AMA1 induces broader allelic recognition and growth inhibition responses in rabbits. *Infect. Immun.*, 2660-2670.
- Remarque, E.J., Faber, B.W., Kocken, CH. M., and Thomas, A.W. (2007), Apical membrane antigen 1: a malaria vaccine candidate in review. *Trends Parasitol.*, 24, 74-84.
- Roux, L., Simon, A.E., Holland, J.J. (1991). Effects of Defective Interfering Viruses on virus replication and pathogenesis in vitro and in vivo. *Adv. Virus Res.*, 40, 181-211.
- Singh M. R., Cattaneo, R., Billeter, M.A. (1999). A recombinant measles virus expressing hepatitis B virus surface antigen induces humoral immune responses in genetically modified mice. *J. Virol.*, 73: 4823-4828.
- Wang, Z.L., Hangartner, L., Cornu, T.I.; Martin, L. R., Zuniga, A., and Billeter, M. (2001). Recombinant measles viruses expressing Heterologous antigens of mumps and simian immunodeficiency viruses. *Vaccine*, 19, 2329-2336.

* cited by examiner

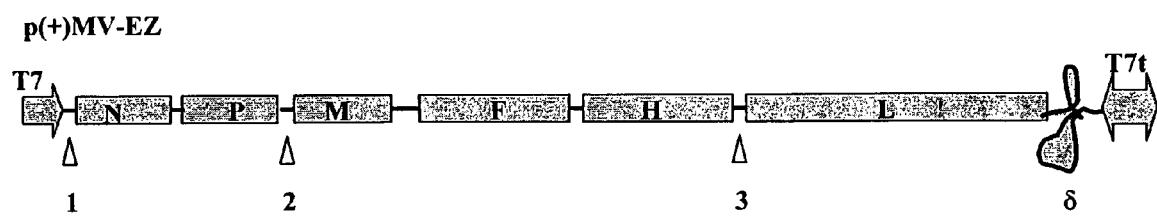
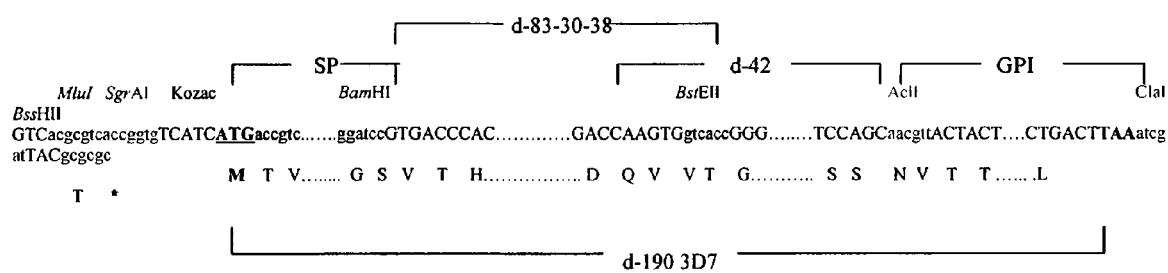


Figure 1

MSP-1 3D7 gene**Figure 2**

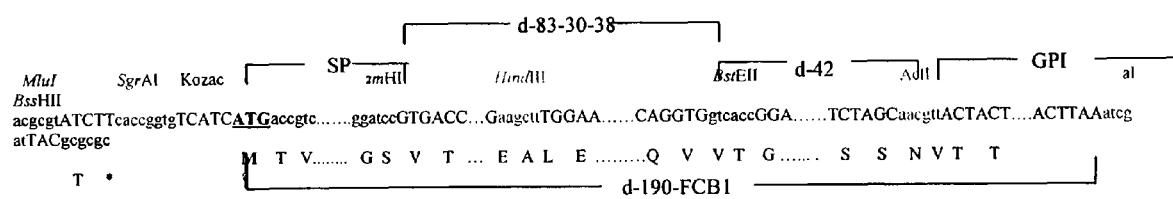


Figure 3

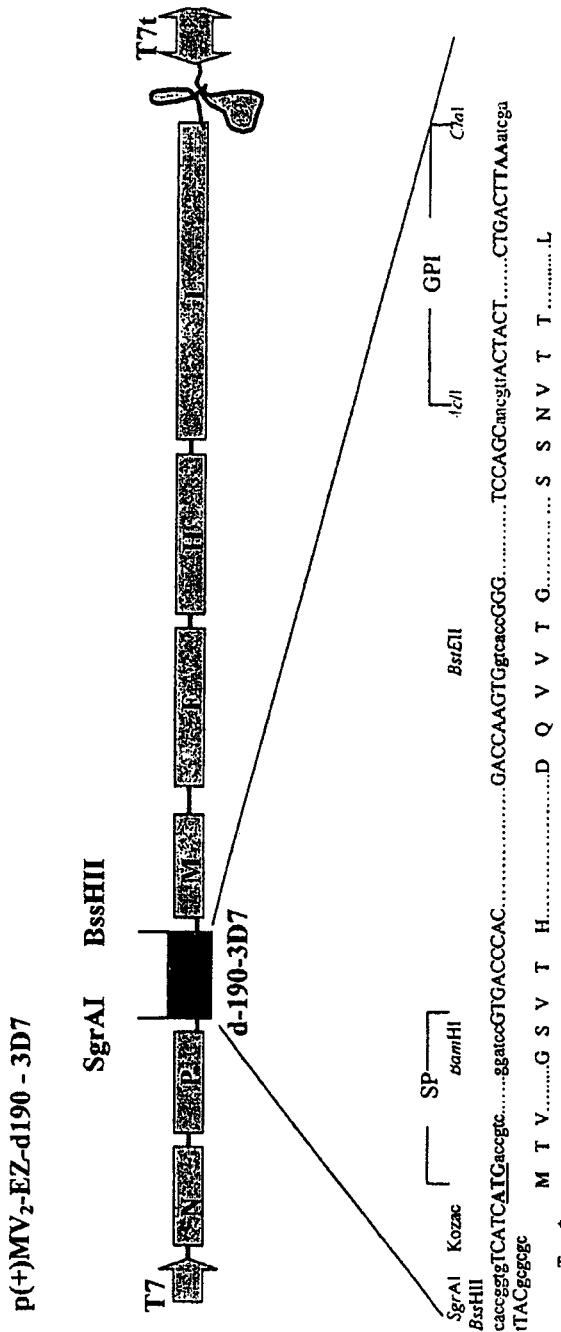


Figure 4

p(+)MV₂-EZ-d190*-3D7

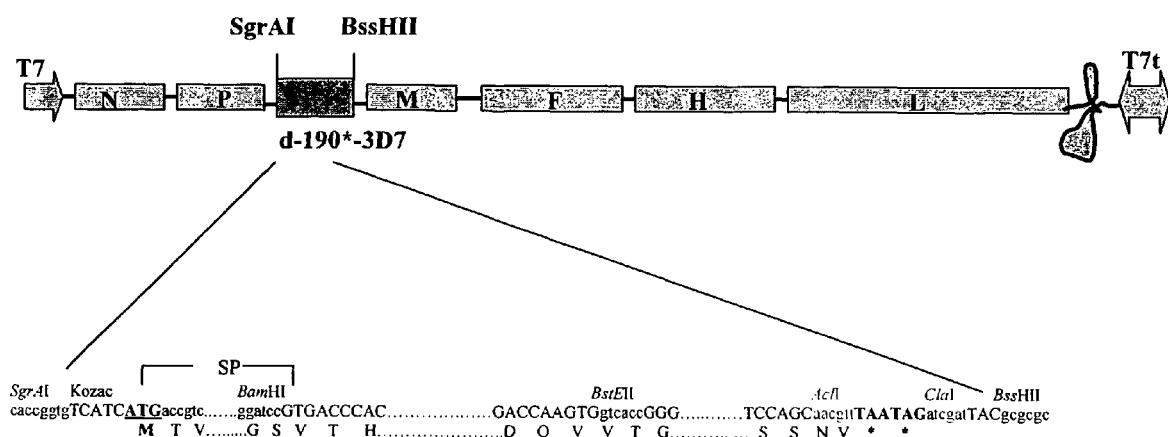


Figure 5

p(+)MV₃-EZ-d190/d190*-3D7

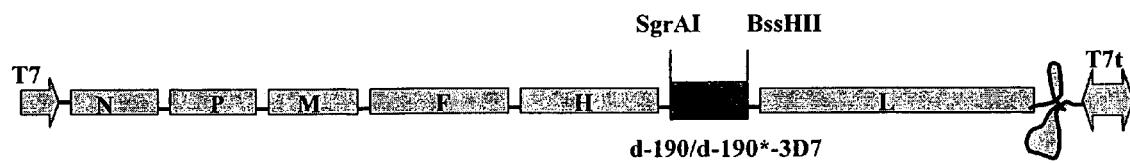


Figure 6

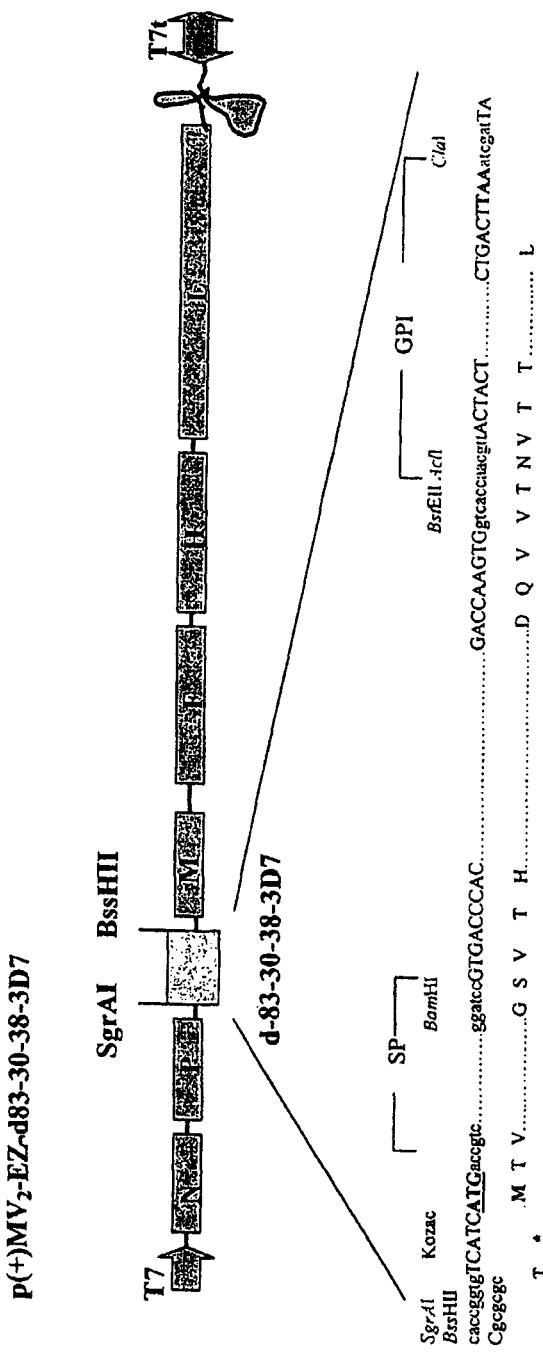


Figure 7

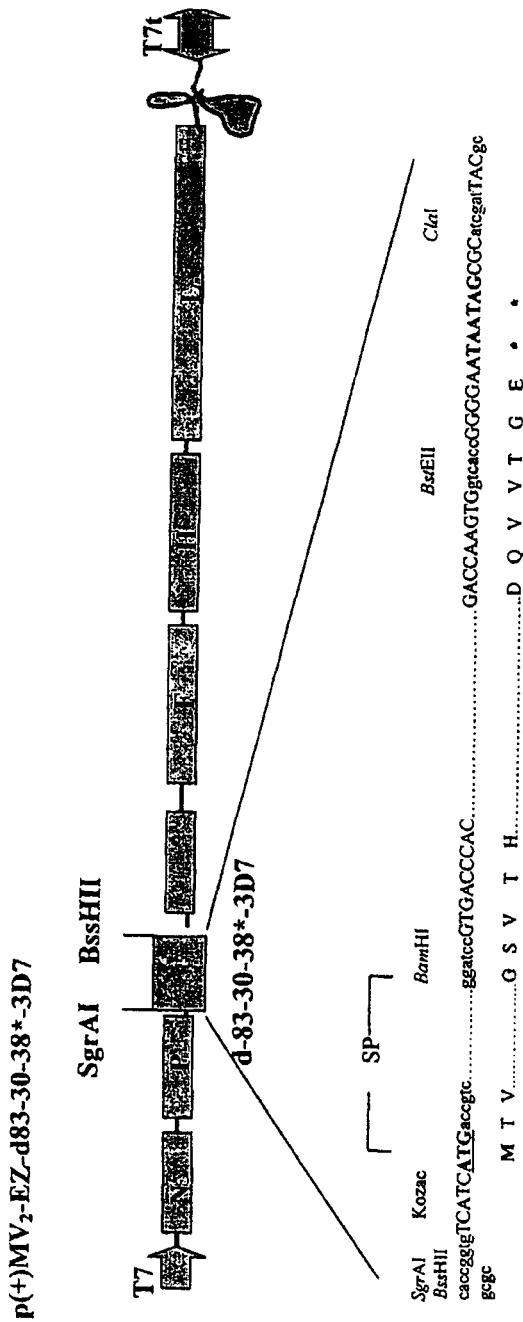


Figure 8

p(+)MV₃-EZ-d-83-30-38/ d-83-30-38*-3D7

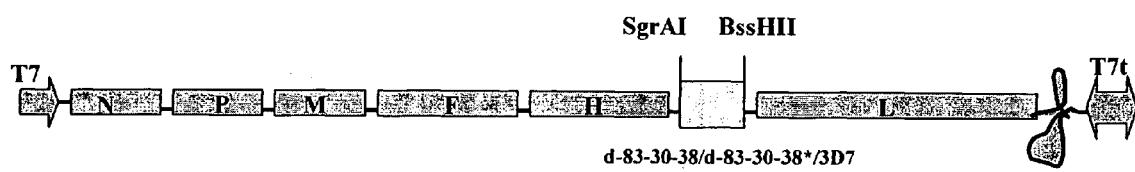
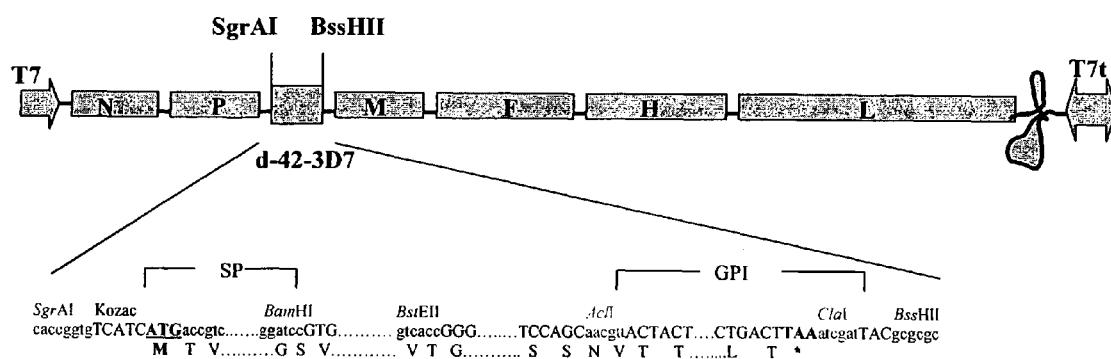


Figure 9

p(+)MV₂-EZ-d-42-3D7**Figure 10**

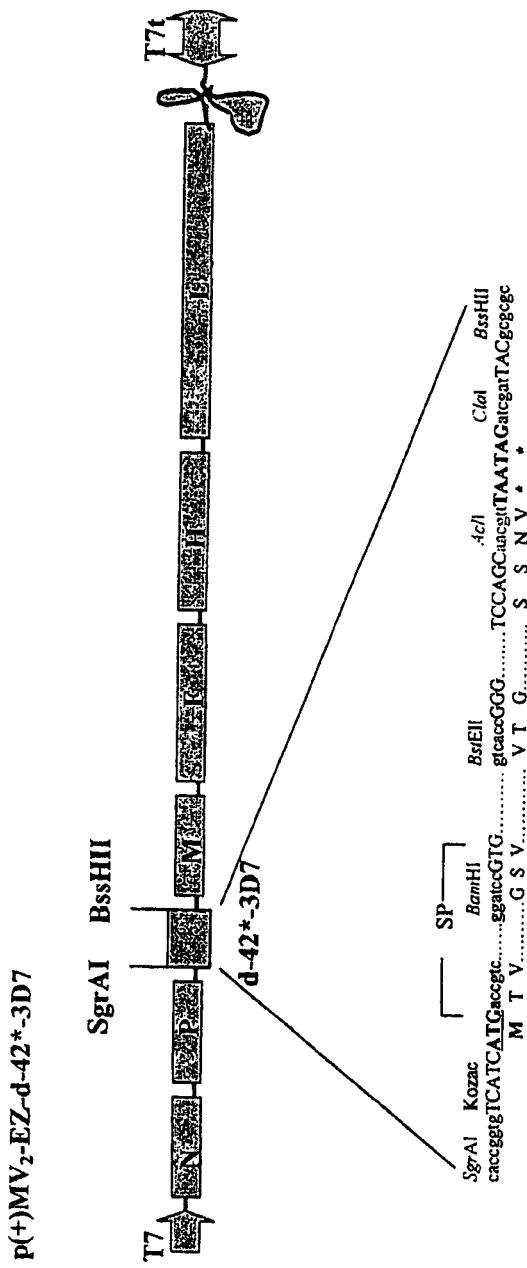


Figure 11

p(+)MV₃-EZ-d-42/d-42*-3D7

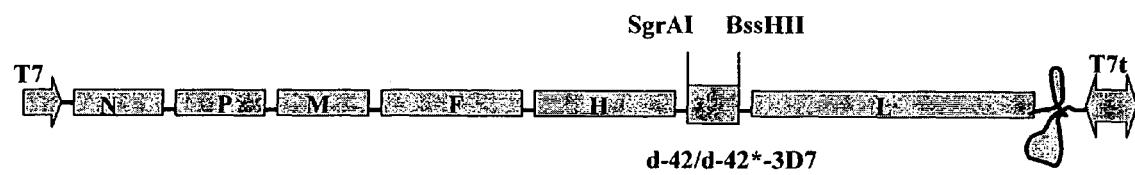


Figure 12

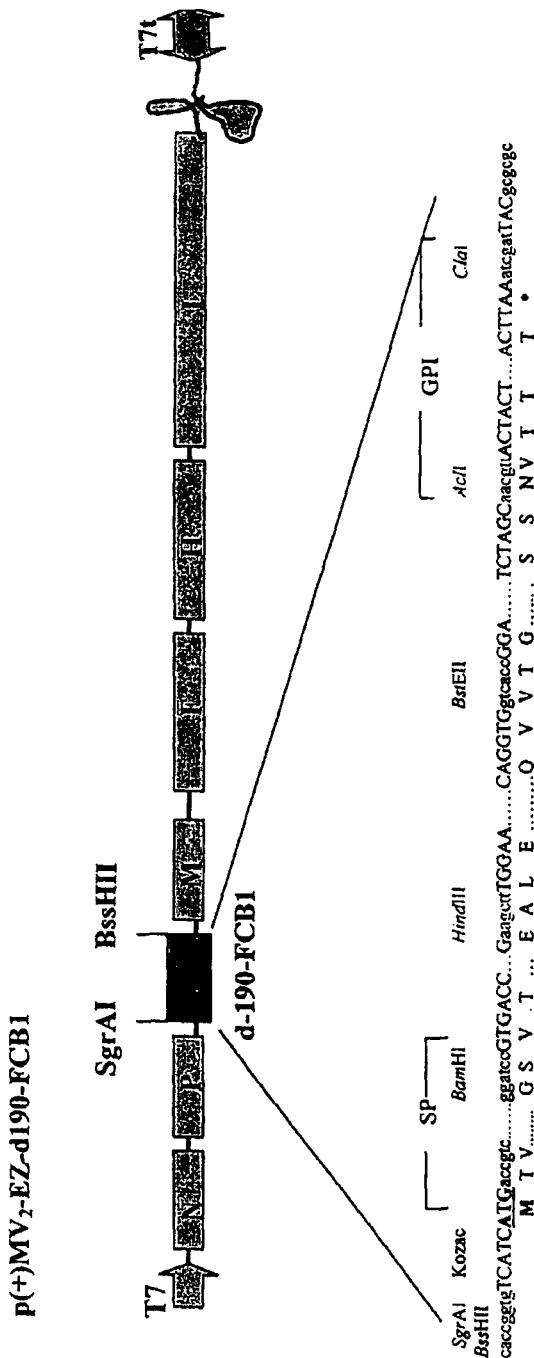


Figure 13

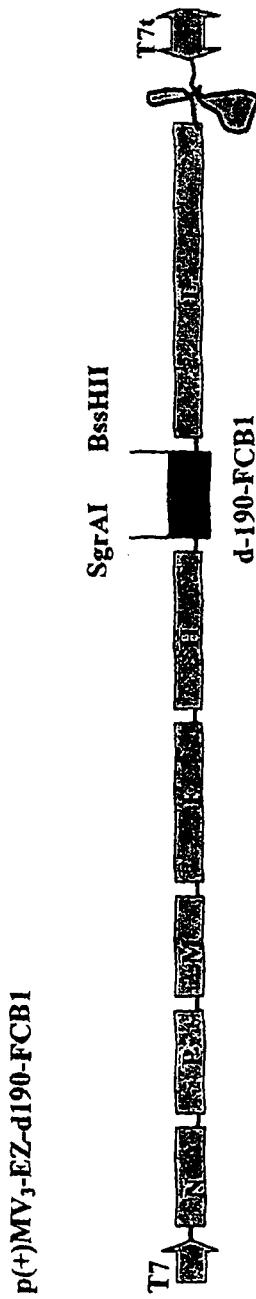


Figure 14

CS gene

BssHII *MspI* *SgrAI* *HindIII* *XbaI*
 acgctATCTTcacgggtTGGAagctGCCACCATG**ATG**GAGGAAACTGGCC.....GTGAACCTCCTGA.....tctagagcgcc
 M R K L A V N S *

Figure 15

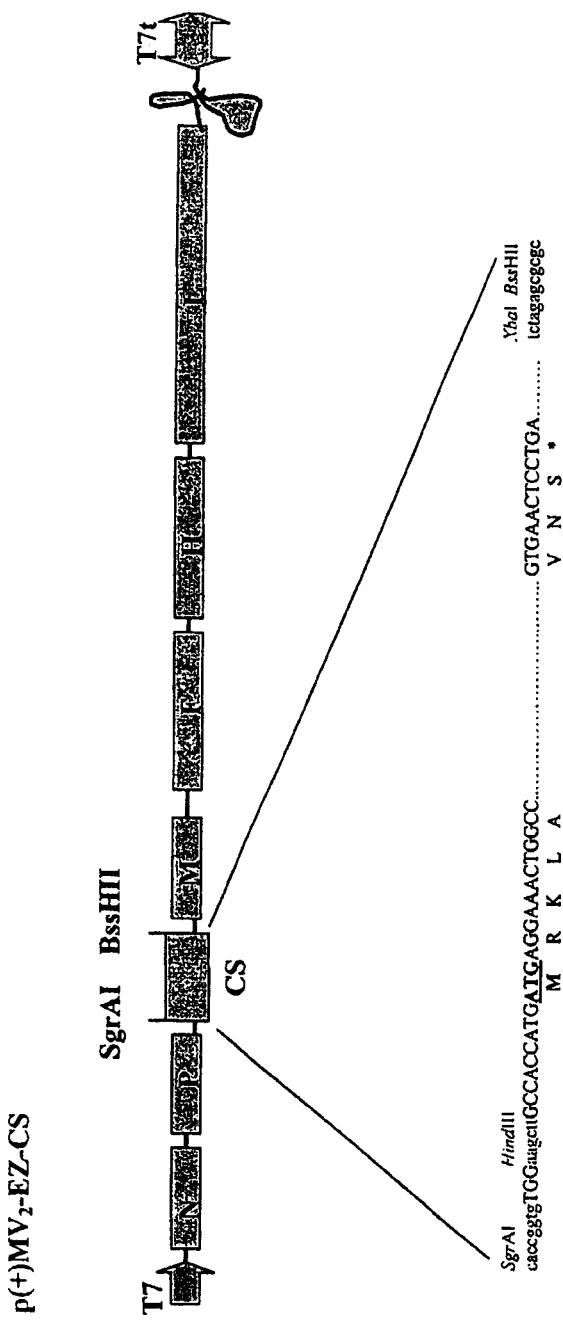


Figure 16

p(+)MV₃-EZ-CS

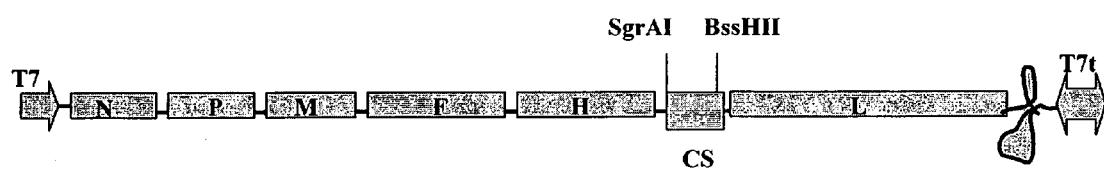


Figure 17

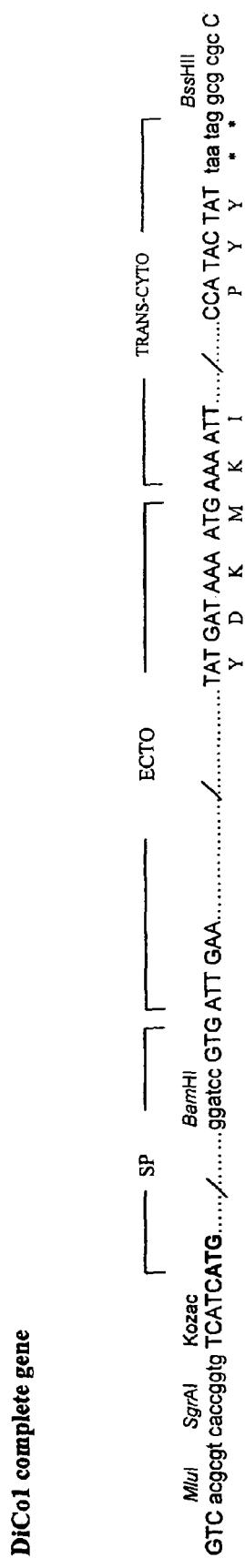


Figure 18



Figure 19

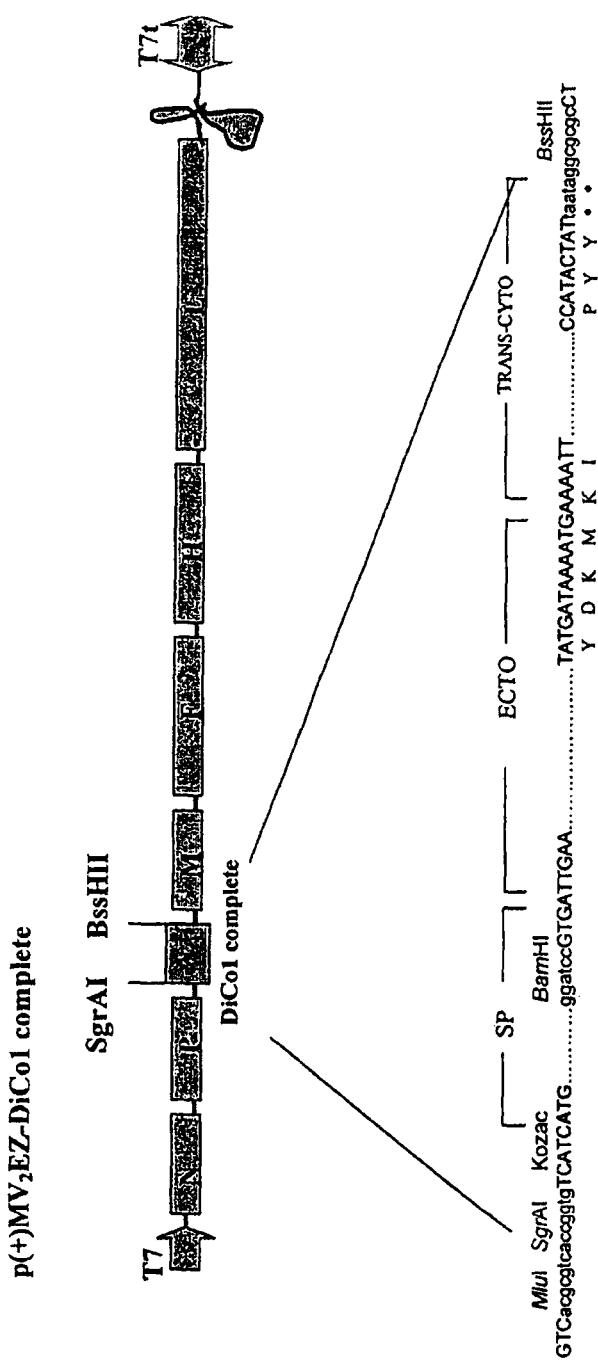


Figure 20

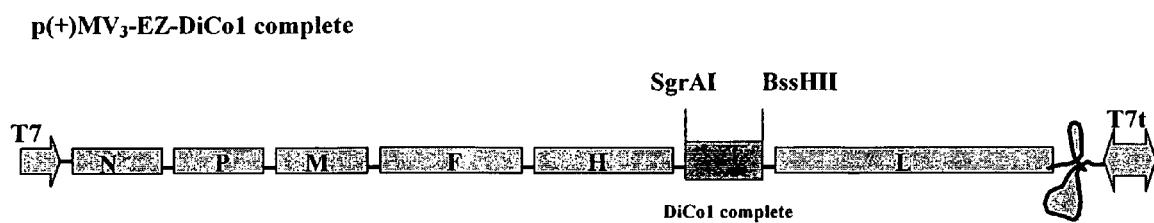


Figure 21

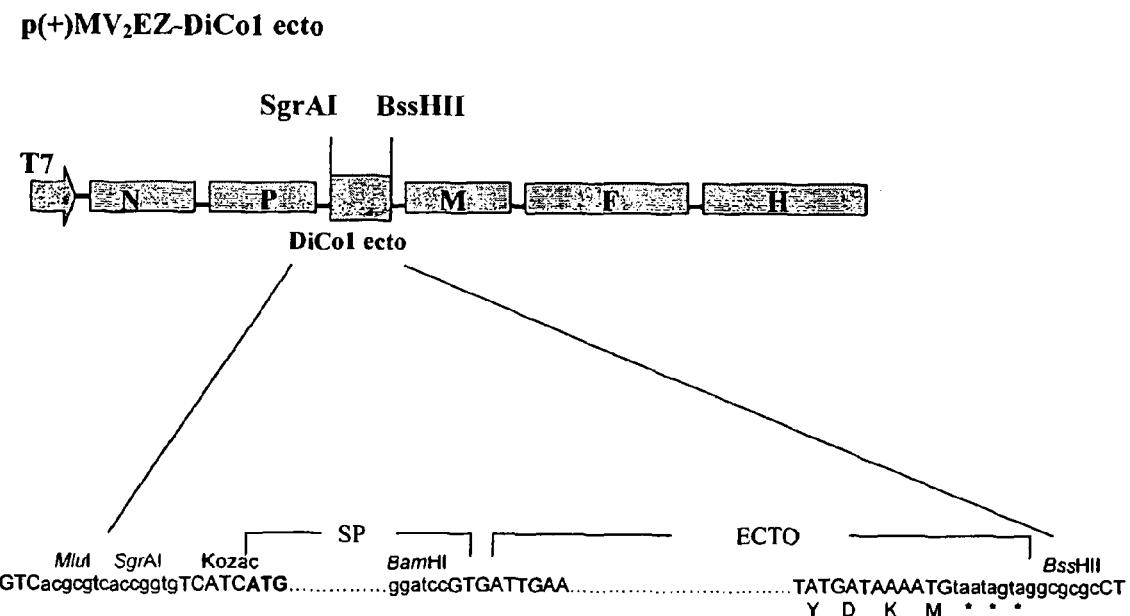


Figure 22

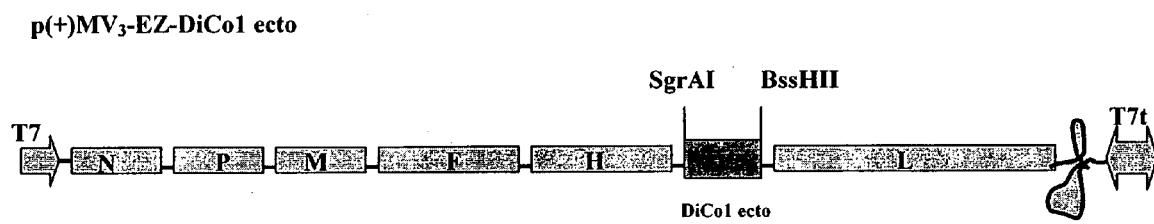


Figure 23

	1	10	20	30	40	50	60	70	80
									90
									100
1	CACCTAAATT	GTAAGCGTTA	ATATTTGTT	AAAATCGCG	TTAAATTTC	GTTAAATCAG	CTCATTTTT		
	AACCAATAGG	CCGAATCGG	CAAATCCCT	100					
101	TATAATCCTA	AAGAATAGAC	CGAGATAGGG	TTGAGTGTG	TTCCAGTTG	GAACAAGAGT	CCACTATTAA		
	AGAACGCTGA	CTCCAACGTC	AAAGGGCGAA	200					
201	AAACCGTCTA	TCAGGGCGAT	GGCCCACTAC	GTGAACCATC	ACCCCTAATCA	AGTTTTTGG	GGTCGAGGTG		
	CCGTAACGCA	CTAAATCGGA	ACCTTAAAGG	300					
301	GAGCCCCCG	TTTAGAGCTT	GACGGGGAAA	CCCggccatt	taggccaTAG	GGCGCTGGCA	AGTGTAGCGG		
	TCACGCTCGG	CGTAACCACC	ACACCCGCCG	400					
401	CGCTTAAATGC	GCCGCTACAG	GGCCGTCCC	ATTCGCCATT	CAGGCTGC	AACTGTTGG	AAGGGCGATC		
	GGTGCAGGCC	TCTTCGCTAT	TACGCCAGCT	500					
501	GGCGAAAGGG	GGATGTGCTG	CAAGGCATT	AGTTGGTA	ACGCCAGGT	TTTCCCAGTC	ACGACGTTGT		
	AAAACGACGG	CCACTGAAATT	Gtaatacgac	600					
601	tcaactataAC	CAAACAAAGT	TGGGTAAGGA	TAGTTCAATC	AATGATCATC	TTCTAGTGCA	CTTAGGATTTC		
	AAGATCCTAT	TATCAGGGAC	AAAGCAGGAA	700					
701	TTAGGGATAT	CTGAGATGGC	CACACTTTA	AGGAGCTTAG	CATTGTTCAA	AAGAAACAAG	GACAAACCAC		
	CCATTACATC	AGGATCCGGT	GGAGCCATCA	800					
801	GAGGAATCAA	ACACATTATT	ATAGTACCAA	TCCCTGGAGA	TTCCCTCAATT	ACCACTCGAT	CCAGACTTCT		
	GGACCGGTG	GTCAGCTTAA	TTGAAACCC	900					
901	GGATGTGAGC	GGGCCCCAAC	TAACAGGGC	ACTAATAGGT	ATATTATCCT	TATTTGTGGA	GTCTCCAGGT		
	CAATTGATTC	AGAGGATCAC	CGATGACCT	1000					
1001	GACCTTAGCA	TAAGGCTGTT	AGAGGTTGTC	CAGAGTGC	AGTCACAAATC	TGGCCTTACC	TTCGCATCAA		
	GAGGTACCA	CATGGAGGAT	GAGGCGGACC	1100					
1101	AATACTTTTC	ACATGATGAT	CCAAATTAGTA	GTGATCAATC	CAGGTTCGGA	TGGTTCGAGA	ACAAGGAAT		
	CTCAGATATT	GAAGTCCGAG	ACCTTGAGGG	1200					
1201	ATTCACATC	ATTCTGGGT	CCATCCTAGC	CCAAATTGG	GTCTTGCTCG	CAAAGGCGGT	TACGGCCCCA		
	GACACGGAG	CTGATTGGGA	GCTAAGAAGG	1300					
1301	TGGATAAAAGT	ACACCCAAAC	AAGAAGGGTA	GTTGGTGAAT	TTAGATTGGA	GAGAAAATGG	TTGGATGTGG		
	TGAGGAACAG	GATTGCCGAG	GACCTCTCCT	1400					
1401	TACGCGGATT	CATGGTCGCT	CTAACCTCTGG	ATATCAAGAG	AACACCCGA	AACAAACCA	GGATTGCTGA		
	ATATGATATGT	GACATGATA	CATATATCGT	1500					
1501	AGAGGCAGGA	TTAGCCAGT	TTATCCTGAC	TATTAAGTTT	GGGATAGAAA	CTATGTATCC	TGCTCTTGG		
	CTGCTGAAT	TTGCTGGTGA	GTATTCACCA	1600					
1601	CTTGACTCCT	TGATGAACCT	TTACCAAGCA	ATGGGGAAA	CTGCACCCCTA	CATGGTAATC	CTGGAGAACT		
	CAATTCAAGAA	CAAGTCAGT	GCAGGATCAT	1700					
1701	ACCCCTGCT	CTGGAGCTAT	GCCATGGGAG	TAGGAGTGA	ACTTGAAAAC	TCCATGGGG	GTTTGAACCT		
	TGGCCGATCT	TACTTTGATC	CAGCATATT	1800					
1801	TAGATTAGGG	CAAGAGATGG	TAAGGAGGT	AGCTGGAAAG	GTCAGTTCCA	CATTGGCATC	TGAACCTCGGT		
	ATCACTGCCG	AGGATGCAAG	GCTTGTTC	1900					
1901	GAGATTGCAA	TGCATACTAC	TGAGGACAAG	ATCAGTAGAG	CGGTTGGACC	CAGACAAGCC	CAAGTATCAT		
	TTCTACACGG	TGATCAAAGT	GAGAATGAGC	2000					
2001	TACCGAGATT	GGGGGGCGA	GAAGATAGGA	GGGTCAAAC	GAGTCGAGGA	GAAGCCAGGG	AGAGCTACAG		
	AGAAACCCGG	CCCACAGAG	CAAGTGTAGC	2100					
2101	GAGAGCTGCC	CATCTTCAA	CCGGCACACC	CCTAGACATT	GACACTGCAT	CGGAGTCCAG	CCAAGATCCG		
	CAGGACAGTC	GAAGGTCA	TGACGCCCTG	2200					
2201	CTTAGGCTG	AAGCCATGGC	AGGAATCTCG	GAAGAACAA	GCTCAGACAC	GGACACCCCT	ATAGTGTACA		
	ATGACAGAAA	TCTTCTAGAC	TAGGTGCGAG	2300					
2301	AGGCCGAGGG	CCAGAACAA	ATCCGCCTAC	CCTCCATCAT	TGTTATAAAA	AACTTAGGAA	CCAGGTCCAC		
	ACAGCCGCCA	CCCCATCAAC	CATCCACTC	2400					
2401	CACGATTGGA	GCCATGGT	GAAGACAGG	CACGCCATGT	AAAAAACGGA	CTGGAATGCA	TCCGGCTCT		
	CAAGGCCGAC	CCCACGGCT	CACTGGCCAT	2500					
2501	CGAGGAAGCT	ATGGCAGCAT	GGTCAGAAAT	ATCAGACAAC	CCAGGACAGG	AGCGAGCCAC	CTGCAGGGAA		
	GAGAAGGCAG	GCAGTTCGGG	TCTCAGAAAA	2600					
2601	CCATGCCCTC	CAGCAATTGG	ATCAACTGAA	GGCGGTGCA	CTCGCATCCG	CGGTCAAGGG	CCTGGAGAGA		
	GGCATGACCA	CGCTGAAACT	TTGGGAATCC	2700					
2701	CCCCAAGAAA	TCTCCAGGCA	TCAGGACTG	GGTTACAGTG	TTATTACGTT	TATGATCACA	GCGGTGAAGC		
	GGTTAAGGG	ATCCAAGATG	CTGACTCTAT	2800					
2801	CATGGTTCAA	TCAGGCCCTG	ATGGTGTAG	CACCCCTCTCA	GGAGGAGACA	ATGAATCTGA	AAACAGCGAT		
	GTGGATATTG	GCAGAACCTGA	TACCGAGGG	2900					
2901	TATGCTATCA	CTGACCCGGG	ATCTGCTCC	ATCTCTATGG	GGTCAGGGC	TTCTGATGTT	GAAACTGCA		
	AAGGAGGGGA	GATCCACGAG	CTCCTGAGAC	3000					
3001	TCCAATCCG	AGGCAACAAAC	TTTCCGAAGC	TTCCCAAAC	TCTCAATGTT	CCTCCGCCCC	CGGACCCCGG		
	TAGGGCCAGC	ACTTCCGGGA	CACCCATTAA	3100					
3101	AAAGGGCACA	GACGCGAGAT	TAGCCTCATT	TGGAACGGAG	ATCGCGTCTT	TATTGACAGG	TGGTGCAC		
	CAATGTGCTC	GAAGATCACC	CTCGGAACCA	3200					
3201	TCAGGGCCAG	GTGCACCTGC	GGGAATGTC	CCCGAGTGTG	TGAGCAATGC	CGCACTGATA	CAGGAGTGG		
	CACCGAATC	TGGTACCCACA	ATCTCCCCGA	3300					
3301	GATCCCCAGAA	TAATGAAGAA	GGGGGAGACT	ATTATGATGA	TGAGCTGTT	TCTGATGTCC	AAGATATTAA		
	AAACGCTTG	GCCAAAATAC	ACGAGGATAA	3400					

Figure 24

3401 TCAGAAGATA ATCTCCAAGC TAGAACATCT GCTGTTATTG AAGGGAGAAG TTGAGTCAAT TAAGAAGCAG
 ATCAACAGGC AAAATATCG CATAATCCACC 3500
 3501 CTGGAAGGAC ACCTCTCAAG CATCATGATC GCCATTCTG GACTTGGAA GGATCCCAC GACCCCACTG
 CAGATGTCGA AATCAATCCC GACTTGAAAC 3600
 3601 CCATCATAGG CAGAGATTCA GGCCGAGCAC TGGCCGAAGT TCTCAAGAAA CCCGTTGCCA GCCGACAAC
 CCAAGGAATG ACAAAATGGAC GGACCAGTTC 3700
 3701 CAGAGGAGAC CTGCTGAAGG AATTTCAGCT AAAGCCGATC GGGAAAAAGA TGAGCTCAGC CGTCGGGTTT
 GTTCCCTGACA CCGGCCCTGC ATCACGCACT 3800
 3801 GAAATCCGCT CCATTATAAA ATCCAGCCGG CTAGAGGAGG ATCGGAAGCG TTACCTGATG ACTCTCCTTG
 ATGATATCAA AGGAGCCAAT GATCTTGCCA 3900
 3901 AGTTCCACCA GATGCTGATG AAGATAATAA TGAAGTAGCT ACAGCTCAAC TTACCTGCCA ACCCCATGCC
 AGTCGACCCA actatgtctac cctccatcat 4000
 4001 ttttataaaa aacttaggaa ccaggtccac acagccgcca gcccattcaAc gcgtATCTTc accggtgATC
 TATacgtacgc ATGagt aaaggagaac 4100
 4101 aacttttcac tggagtgtc ccaattcttg ttgaattttaa tggtgtatgtt aatggcaca aatttctgt
 cagtggagag ggtgaagggt atgcaacata 4200
 4201 cgaaaaactt acccttaaat ttatttgac tactggaaaa ctaccgttca catggccaac acttgtcaact
 actttcacat atgggttca atgctttca 4300
 4301 agatccccac atccatgaa acggcatgac ttttcaaga gtgccccatgcc cgaagggtac gtacaggaaa
 gaactatatt ttccaaagat gacggaaat 4400
 4401 acaagacacg tgctgaagtc aagtttgaag gtgataccct tgtaataga atcgagttaa aaggatttga
 ttttaagaa gatggaaaca ttctggaca 4500
 4501 caaatttggaa tacaactata actcacacaa tgtatcacatc atggcagaca aacaaaagaa tggaaatcaga
 gttacttca aatttagaca caacattgaa 4600
 4601 gatggaaagcg ttcaacttgc agaccattat caacaaaata ctccaaattgg cgatggccct gtcctttac
 cagacaacca ttacctgttc acacatctg 4700
 4701 ccctttcgaa agatccccac gaaaagagag accacatggt ccttcttgag tttgtaacag ctgtggat
 tacacatggc atggatgaaac tatacaaata 4800
 4801 gtgagcgcgc aegcgtacg ttcgcgcgtat aatactgtAC AACCTAAATC CATCATAAAA AACTTAGGAG
 CAAAGTGATT GCCTCCCAAG TTCCACAAATG 4900
 4901 ACAGAGATCTC ACGACTTGA CAAGTCGGCA TGGGACATCA AAGGGTCGAT CGCTCCGATA CAACCCACCA
 CCTACAGTGA TGGCAGGCTG GTGGCCCG 5000
 5001 TCAGACTCAT AGATCTGTG CTAGGGCACA GGAAGGATGA ATGCTTATG TACATGTTTC TGCTGGGGT
 TGTTGAGGAC AGGGATTCCC TAGGGCCTCC 5100
 5101 AATCGGGCGA GCATTTGGT CCCTGCCTT AGGTGTTGGC AGATCCACAG CAAAGCCGA AAAACTCCTC
 AAAGAGGCCA CTGAGCTTGA CATAATTGTT 5200
 5201 AGACGTACAG CAGGGCTCAA TGAAAAACTG GTGTTCTACA ACAACACCCC ACTAACTCTC CTCACACCTT
 GGAGAAAGGT CCTAAACACA GGGAGTGTCT 5300
 5301 TCAACGCAAA CCAAGTGTGC AATGCGGTAA ATCTGATACC GCTCGATACC CCGCAGAGGT TCCGTGTTG
 TTATATGAGC ATCACCCGTC TTTCGGATAA 5400
 5401 CGGGTATTAC ACCGTTCTTA GAAGAATGCT GGAATTCAAGA TCGGTCAATG CAGTGGCCTT CAACCTGCTG
 GTGACCCCTTA GGATTGACAA GGGATAGGC 5500
 5501 CCTGGGAAGA TCATCGACAA TACAGACCAA CTTCCGTAGG CAACATTTAT AGTCCACATC GGGAACTTCA
 GGAGAAAGAA GAGTGAAGTC TACTCTGCCG 5600
 5601 ATTATTGCAA AATGAAAATC GAAAAGATGG GCCTGGTTT TGCACTTGGT GGGATAGGGG GCACCAAGTCT
 TCACATTAGA AGCACAGGC AAATGAGCAA 5700
 5701 GACTCTCAAT GCACAACTCG GTTCAAGAA GACCTTATGT TACCCGCTGA TGGATATCAA TGAAGACCTT
 AATCGATTAC TCTGGAGGAG CAGATGCAAG 5800
 5801 ATAGTAAGAA TCCAGGCACT TTGAGGCA TCAAGTTCTC AAGAATTCCG CATTACGAC GACGTGATCA
 TAAATGATGA CCAAGGACTA TTCAAGTTG 5900
 5901 TGTAGACCGT AGTGGCCAGC AATGCCCCAA AACGACCCCC CTCACAATGA CACCCAGAAG GCCCGGACAA
 AAAAGCCCCC TCCGAAAGAC TCCACGGACC 6000
 6001 AAGCGAGAGG CCAGGCCAGA GCCGACGGCA AGCGCGAACCA CCAGGCCGCC CCAGCACAGA ACAGCCCTGA
 CACAAGGCCA CCACCGAGCCA CCCCAATCTG 6100
 6101 CATCTCTCTC GTGGGACCCC CGAGGACCAA CCCCCAAGGC TGCCCCCGAT CCAAAACCAAC AACCGCATCC
 CCACCAACCC CGGGAAAGAA ACCCCAGCA 6200
 6201 ATTGGAAAGGC CCCTCCCCCT CTTCTCAAC ACAAGAAACTC CACAACCGAA CCGCACAAGC GACCGAGGTG
 ACCCAACCGC AGGCATCCGA CTCCCTAGAC 6300
 6301 AGATCCTCTC TCCCCGGCAA ACTAAACAAA ACTTAGGGCC AAGGAACATA CACACCCAAAC AGAACCCAGA
 CCCCCGGCCCA CGGGCCCGCG CCCCAACCC 6400
 6401 CCGACAACCA GAGGGAGCCC CCAACCAATC CGCGCGGCTC CCCCCGGTGCC CACAGGCAGG GACACCAACC
 CCCGAACAGA CCCAGCACCC AACCATCGAC 6500
 6501 AATCCAAGAC GGGGGGGCCC CCCCCAAAAAA AAGCCCCCAG GGGCCGACAG CCAGCACCGC GAGGAAGGCC
 ACCCAACCCA CACACGACCA CGGAACCAA 6600
 6601 ACCAGAACCC AGACCACCTT GGGCACCAG CTCCCAAGACT CGGCCATCAC CCCCGAGAAA GGAAAGGCCA
 CAACCCCGCGC ACCCCAGCCC CGATCCGGCG 6700
 6701 GGGAGCCACC CAACCCGAAC CAGCACCCAA GAGCGATCCC CGAAGGACCC CGGAACCGCA AAGGACATCA
 GTATCCCACCA GCCTCTCCAA GTCCCCCGGT 6800

Figure 24 (Contd..)

6801	CTCCCTCCTCT	TCTCGAAGGG	ACCAAAAGAT	CAATCCACCA	CACCCGACGA	CACTCAACTC	CCCACCCCTA
	AAGGAGACAC	CGGGAATCCC	AGAATCAAGA	6900			
6901	CTCATCCAAT	GTCCATCATG	GGTCTCAAGG	TGAACGTCTC	TGCCATATTG	ATGGCAGTAC	TGTTAACCTCT
	CCAAACACCC	ACCGGTCAAA	TCCATTGGGG	7000			
7001	CAATCTCTCT	AAGATAGGGG	TGGTAGGAAT	AGGAAGTCCA	AGCTACAAAG	TTATGACTCG	TTCCAGGCCAT
	CAATCATTAG	TCATAAAATT	AAATGCCAAT	7100			
7101	ATAACTCTCC	TCATAACTG	CACCGAGGTA	GAGATTGCAG	AATACAGGAG	ACTACTGAGA	ACAGTTTTGG
	AACCAATTAG	AGATGCACTT	AATGCAATGA	7200			
7201	CCCAGAATAT	AAGACCGGTT	CAGAGTGTAG	CTTCAAGTAG	GAGACACAAG	AGATTTGCGG	GAGTAGTCCT
	GGCAGGTGCG	GCCCTAGGCG	TTGCCACAGC	7300			
7301	TGCTCAGATA	ACGGCCGGG	TTGCACTTCA	CCAGTCCATG	CTGAACCTCTC	AAGCCATCGA	CAATCTGAGA
	GGCAGCTCG	AAACTACTAA	TCAGGCAATT	7400			
7401	GAGGCAATCA	GACAAGCAGG	GCAGGAGATG	ATATTGGCTG	TTCAAGGGTGT	CCAAGACTAC	ATCAATAATG
	AGCTGATAACC	GTCTATGNA	CAACTATCTT	7500			
7501	GTGATTTAAT	CGGCCAGAAG	CTCGGGCTCA	AATTGCTAG	ATACTATACA	GAAATCCTGT	CATTATTTGG
	CCCCAGTTA	CGGGACCCCA	TATCTGCGGA	7600			
7601	GATATCTATC	CAGGCTTGA	GCTATCGCCT	TGGAGGAGAC	ATCAATAAGG	TGTTAGAAA	GCTCGGATAC
	AGTGGAGGTG	ATTACTGCG	CATCTTAGAG	7700			
7701	AGCAGAGGAA	TAAGGCCCG	GATAACTCAC	GTCGACACAG	AGTCCTACTT	CATTGTCCTC	AGTATAGCCT
	ATCCGACGCT	GTCCGAGATT	AAGGGGGTGA	7800			
7801	TTGTCCACCG	GCTAGAGGGG	GTCTCGTACA	ACATAGGCTC	TCAAGAGTGG	TATACCACTG	TGCCCAAGTA
	TGTTCCAACC	CAAGGGTAC	TTATCTCGAA	7900			
7901	TTTGATGAG	TCATCGTGA	CTTTCATGCC	AGAGGGGACT	GTGTGCAGCC	AAAATGCCTT	GTACCCGATG
	AGTCCTCTGC	TCCAAGAATG	CCTCCGGGGG	8000			
8001	TACACCAAGT	CCTGTGCTCG	TAACACTCGA	TCCGGTCTT	TTGGGAACCG	GTTCATTTA	TCACAAGGGA
	ACCTAATAGC	CAATTGTCGA	TCAATCCTTT	8100			
8101	GCAAGTGTAA	CACAACAGGA	ACGATCATT	ATCAAGACCC	TGACAAGATC	CTAACATACA	TTGCTGCCGA
	TCACTGCCCG	GTAGTCGAGG	TGAACGGCCT	8200			
8201	GACCATCCAA	GTCGGGAGCA	GGAGGTATCC	AGACGCTGTG	TACTTGCACA	GAATTGACCT	CGGTCTCC
	ATATCATGG	AGAGGTTGGA	CGTAGGGACA	8300			
8301	AATCTGGGA	ATGCAATTGC	TAAGTGGAG	GATGCCAAGG	AATTGTTGA	GTCATCGGAC	CAGATATTGA
	GGAGTATGAA	AGGTTTATCG	AGCACTAGCA	8400			
8401	TAGTCTACAT	CCTGATTGCA	GTGTGTCTTG	GAGGGTGAT	AGGGATCCCC	GCTTTAATAT	GTTGCTGCAG
	GGGGCGTTGT	AAACAAAAAGG	GAGAACAACT	8500			
8501	TGGTATGTCA	AGACCGCC	TAAGCCTGA	TCTTACGGGA	ACATCAAAT	CCTATGTAAG	GTCGCTCTGA
	TCCCTACAA	CTCTTGAAAC	ACAAATGTCT	8600			
8601	CACAAGTCTC	CTCTTCGTCA	TCAACCAACC	ACCGCACCCA	GCATCAAGCC	CACCTGAAAT	TATCTCCGGC
	TTCCCTCTGG	CCGAACAAATA	TCGGTAGTTA	8700			
8701	ATTAACAACTT	AGGGTGAAG	ATCATCCACA	ATGTCAACCAC	AAAGAGACCG	GATAAAATGCC	TTCTACAAAG
	ATAACCCCCA	TCCCAAGGGA	AGTAGGATAG	8800			
8801	TCATTAATGAC	AGAACATCTT	ATGATTGATA	GACCTTATGT	TTTGCCTGGCT	GTTCTGTTG	TCATGTTCT
	GAGCTTGATC	GGGTTGCTAG	CCATTGCGAG	8900			
8901	CATTAGACTT	CATCGGGCAG	CCATCTACAC	CGCAGAGATC	CATAAAAGCC	TCAGCACCAA	TCTAGATGTA
	ACTAACTCAA	TCGAGCATCA	GGTCAGGAC	9000			
9001	GTGCTGACAC	CACTCTCAA	AATCATCGGT	GATGAAGTGG	GCCTGAGGAC	ACCTCAGAGA	TTCACTGACC
	TAAGTAAATT	CATCTCTGAC	AAAGATTAAAT	9100			
9101	TCCTTAATCTC	GGATAGGGG	TACGACTTCA	GAGATCTCAC	TTGGTGTATC	AACCCGCCAG	AGAGAATCAA
	ATTGGATTAT	GATCAATACT	GTGCAAGATGT	9200			
9201	GGCTGCTGAA	GAGCTCATGA	ATGCATTGGT	GAACCTCAACT	CTACTGGAGA	CCAGAACAAAC	CAATCAGTTC
	CTAGCTGTCT	CAAAGGGAA	CTGTCAGGG	9300			
9301	CCCACTACAA	TCAGAGGTCA	ATTCTCAAAC	ATGTCGCTGT	CCCTGTAGA	CTTGTATTTA	GGTCGAGGTT
	ACAATGTGTC	ATCTATAGTC	ACTATGACAT	9400			
9401	CCCAGGGAA	GTATGGGG	ACTTACCTAG	TGGAAAAGCC	TAATCTGAGC	AGCAAAAGGT	CAGAGTTGTC
	ACAACTGAGC	ATGTACCGAG	TGTTGAGCT	9500			
9501	AGGTGTTATC	AGAAATCCGG	GTTTGGGGC	TCCGGTGTTC	CATATGACAA	ACTATCTTGA	GCAACCAGTC
	AGTAATGATC	TCAGCAACTG	TATGGTGGCT	9600			
9601	TTGGGGAGC	TCAAACCTCG	AGCCCTTGT	CACGGGAAG	ATTCTATCAC	AATTCCCTAT	CAGGGATCAG
	GGAAAGGTGT	CAGCTTCCAG	CTCTCAAGC	9700			
9701	TAGGTGTCTG	GAATCCCCA	ACCGACATGC	AATCCTGGGT	CCCCTTATCA	ACGGATGATC	CAGTGTAGA
	CAGGCTTAC	CTCTCATCTC	ACAGAGGTGT	9800			
9801	TATCGCTGAC	AATCAAGCAA	AATGGGCTGT	CCCGACAAAC	CGAACAGATG	ACAAGTTGCG	AATGGGAGACA
	TGCTTCCAAC	AGGCCTGTAA	GGGAAAATC	9900			
9901	CAAGCACTCT	CGGAGAATCC	CGAGTGGGCA	CCATTGAAGG	ATAACAGGAT	TCCTTCATAC	GGGGTCTTGT
	CTGTTGATCT	GAGTCTGACA	GTGAGCTTA	10000			
10001	AAATCAAAAT	TGCTTCGGGA	TTCCGGCCAT	TGATCACACA	CGGTTCAAGG	ATGGACCTAT	ACAAATCAA
	CCACAAACAAAT	GTGTATTGGC	TGACTATCCC	10100			
10101	GCCAATGAAG	AACTAGCCT	TAGGTGTAAT	CAACACATTG	GAGTGGATAC	CGAGATTCAA	GGTTAGTCCC
	TACCTCTTCA	CTGTCCCAAT	TAAGGAAGCA	10200			

Figure 24 (contd..)

10201 GGCGAAGACT GCCATGCCCT AACATACCTA CCTGCGGAGG TGGATGGTGA TGTCAAACTC AGTTCCAATC
 TGGTGATTCT ACCTGGTCAA GATCTCCAAT 10300
 10301 ATGTTTGCG AACCTACCGAT ACTTCCAGGG TTGAACATGC TGTGGTTAT TACCTTACA GCCCAGGCCG
 CTCATTTCT TACTTTTATC CTTTTAGGTT 10400
 10401 GCCTATAAAG GGGGTCCTCA TCGAATTACA AGTGAATGC TTCACATGGG ACCAAAAACT CTGGTGCCGT
 CACTTCTGTG TGCTTGCGGA CTCAGAAATCT 10500
 10501 GGTGGACATA TCACTCACTC TGGGATGGTG GGCATGGAG TCAGCTGCAC AGTCACCCGG GAAGATGGAA
 CCAATCGAG ATAGGGCTG TAGTGAACCA 10600
 10601 ATCACATGAT GTCAACCCAGA CATCAGGCAT ACCCACTAGT GTGAAATAGA CATCAGAATT AAGAAAAACG
 TAGGGTCCAA GTGGTTCCCC GTTATGGACT 10700
 10701 CGCTATCTGT CAACCAAGATC TTATACCCCTG AAGTTCACCT AGATAGCCCG ATAGTTACCA ATAAGATAGT
 AGCCATCTG GAGTATGCTC GAGTCCCTCA 10800
 10801 CGCTTACAGC CTGGAGGACC CTACACTGTG TCAGAACATC AAGCACCAGC TAAAAAACGG ATTTCCAAC
 CAAATGATTA TAAACAAATGT GGAAGTTGGG 10900
 10901 AATGTCATCA AGTCAAGCT TAGGAGTTAT CCGGCCCCACT CTCATATTCC ATATCCAAT TGTAATCAGG
 ATTATTATTA CATAGAAGAC AAAGACTCAA 11000
 11001 CGAGGAAGAT CCGTGAACTC CTCAAAAAGG GGAATTGCGT GTACTCCAAA GTCAGTGATA AGGTTTCCA
 ATGCTTAAGG GACACTAATC CACGGCTTGG 11100
 11101 CCTAGGCTCG GAATTGAGGG AGGACATCAA GGAGAAAGTT ATTAACCTGG GAGTTACAT GCACAGCTCC
 CAGTGGTTTG AGCCCTTTCT GTTTGGTTT 11200
 11201 ACAGTCAGA CTGAGATGAG GTCAGTGATT AAATCACAAA CCCATACCTG CCATAGGAGG AGACACACAC
 CTGTATTCTT CACTGGTAGT TCAGTTGAGT 11300
 11301 TGCTAATCTC TCGTGAACCTT GTTGCTATAA TCAGTAAAGA GTCTCAACAT GTATATTACC TGACATTGGA
 ACTCGTTTG ATGTATTGTG ATGTATAGA 11400
 11401 GGGGAGGTT ATGACAGAGA CCCCTATGAC TATTGATGCT AGGTATACAG AGCTTCTAGG AAGAGTCAGA
 TACATGTGGA AACTGATAGA TGGTTCTTC 11500
 11501 CCTGCACTCG GGAATCCAC TTATCAAATT GTACCAATGC TGGAGCCTCT TTCACTTGCT TACCTGCAGC
 TGAGGGATAT AACAGTAGAA CTCAGAGGTG 11600
 11601 CTTTCCTTAA CCACTGCTT ACTGAAATAC ATGATGTTCT TGACCAAAC GGGTTTCTG ATGAAGGTAC
 TTATCATGAG TTAATTGAAG CTCTAGATTA 11700
 11701 CATTTCATA ACTGTGACAA TACATCTGAC AGGGGAGATT TTCTCATTTC TCAGAAGTTT CGGCCACCCC
 AGACTTGAAG CAGTACAGG TGCTGAAAAT 11800
 11801 GTTAAAGAT ACATGAATCA GCCTAAAGTC ATTGTGTATG AGACTCTGAT GAAAGGTAT GGCATATTIT
 GTGAAATCAT AATCAACGGC TATCGTGACA 11900
 11901 GGCACGGAGG CAGTTGCCA CGCCTGACCC TCCCCCTGCA TGCTGCAGAC ACAATCCGGA ATGCTCAAGC
 TTCAAGGTGAA GGGTTAACAC ATGAGCAGTG 12000
 12001 CGTTGATACT TGGAAATCTT TTGCTGGAGT GAAATTGGC TGCTTATGC CTCTTAGCCT GGATAGTGAT
 CTGACAATGT ACCTAAAGGA CAAGGCACTT 12100
 12101 GCTGCTCTC AAAGGGATG GGATTCAAGT TACCCGAAAG AGTTCCGTGCG TTACGACCTC CCCAAGGGAA
 CCGGGTCAC GAGGCTGTGTA GATCTTTCC 12200
 12201 TTAATGATTC GAGCTTGAC CCATATGATG TGATAATGTA TGTTGTAAGT GGAGCTTACCC TCCATGACCC
 TGAGTTCAAC CTGTCTTACA GCCTGAAAGA 12300
 12301 AAAGGAGATC AAGGAAACAG GTAGACTTT TGCTAAAATG ACTTACAAA TGAGGGCATG CCAAGTGATT
 GCTGAAATTC TAATCTCAA CGGGATTGGC 12400
 12401 AAATATTTA AGGACAATGG GATGCCAAG GATGAGCACG ATTTGACTAA GGCACCCAC ACTCTAGCTG
 TCTCAGGAGT CCCCAAAGAT CTAAAGAAA 12500
 12501 GTCACAGGGG GGGGCCAGTC TTAACCAACCT ACTCCCGAAG CCCAGTCCAC ACAAGTACCA GGAACGTGAG
 AGCAGCAAA GGGTTTATAG GGTTCCCTCA 12600
 12601 AGTAATTCGG CAGGACAAAG ACACTGATCA TCCGGAGAAT ATGGAAGCTT ACGAGACAGT CAGTGCATT
 ATCACGACTG ATCTCAAGAA GTACTGCCTT 12700
 12701 AATTGGAGAT ATGAGACCAT CACCTGTTT GCACAGAGGC TAAATGAGAT TTACGGATTG CCCTCATTTC
 TCCAGTGGCT GCATAAGAGG CTTGAGACCT 12800
 12801 CTGTCCTGTA TGTAAAGTGAC CCTCATTGCC CCCCCGACCT TGACGCCAT ATCCCGTTAT ATAAAGTCCC
 CAATGATCAA ATCTTCATTA AGTACCCCTAT 12900
 12901 GGGAGGTATA GAAGGGTATT GTCAAGAGCT GTGGACCATC AGCACCATTC CCTATCTATA CCTGGCTGCT
 TATGAGAGGC GAGTAAGGAT TGCTCTGTTA 13000
 13001 GTGCAAGGGG ACAATCAGAC CATAGCCGTAA ACAAAAAGGG TACCCAGCAC ATGGCCCTAC AACCTTAAGA
 AACGGGAAGC TGCTAGAGTA ACTAGAGATT 13100
 13101 ACTTTGTAAT TCTTAGGCAA AGGCTACATG ATATTGGCCA TCACCTCAAG GCAAATGAGA CAATTGTTTC
 ATCACATTTC TTGTCCTATT CAAAGGAAT 13200
 13201 ATATTATGAT GGGCTACTTG TGCCCAATC ACTCAAGAGC ATCGCAAGAT GTGTATTCTG GTCAGAGACT
 ATAGTTGATG AAAACAAGGGC AGCATGCACT 13300
 13301 AATATTGCTA CAACATGGC TAAAGCCTAC GAGAGAGGTT ATGACCGTTA CCTTGCATAT TCCCTGAACG
 TCTCTAAAGT GATACAGCAA ATTCTGATCT 13400
 13401 CTCTTGCTT CACAATCAAT TCAACCATGA CCCGGATGT AGTCATACCC CTCCTCACAA ACAACGACCT
 CTTAATAAGG ATGGCACTGT TGCCGCTCC 13500
 13501 TATTGGGGGG ATGAATTATC TGAATATGAG CAGGCTGTTT GTCAGAAACA TCGGTGATCC AGTAACATCA
 TCAATTGCTG ATCTCAAGAG AATGATTCTC 13600
 13601 GCCTCACTAA TGCTGAAGA GACCCCTCAT CAAGTAATGA CACAACAAACC GGGGGACTCT TCATTCCCTAG

Figure 24 (Contd..)

ACTGGGCTAG CGACCCATTAC TCAGCAAATC 13700
 13701 TTGTATGTGT CCAGAGCATC ACTAGACTCC TCAAGAACAT AACTGCAAGG TTTGTCCTGA TCCATAGTCC
 AAACCCAATG TAAAGGAT TATTCATGA 13800
 13801 TGACAGTAAA GAAGAGGACG AGGGACTGGC GGCATTCTC ATGGACAGGC ATATTATAGT ACCTAGGGCA
 GCTCATGAAA TCCGGATCA TAGTGTACA 13900
 13901 GGGCAAGAG AGTCTATGTC AGGCATGCTG GATACCACAA AAGGCTTGAT TCGAGCCAGC ATGAGGAAGG
 GGGGGTTAAC CTCTCGAGTG ATAACCAGAT 14000
 14001 TGTCATTAA TGACTATGAA CAATTCAAGAG CAGGGATGGT GCTATTGACA GGAAGAAAGA GAAATGTCC
 CATTGACAAA GAGTCATGTT CAGTGTGAGCT 14100
 14101 GGCGAGAGCT CTAAGAACCC ATATGTGGC GAGGGTAGCT CGAGGACGGC CTATTTACGG CCTTGAGGTC
 CCTGTGATGAC TAGAATCTAT CGCGGAGGCCAC 14200
 14201 CTTATTGGC GTCTGAGAC ATGTGTGATC TCGGAGTGTG CATCAGTCAA CTACGGATGG TTTTTGTCC
 CCTCGGGTTG CCAACTGGAT GATATTGACA 14300
 14301 AGGAAACATC ATCCTTGAGA GTCCCATATA TTGGTTCTAC CACTGATGAG AGAACAGACA TGAAGCTTGC
 CTTCGTAAGA GCCCGAACGTC GATCCTTGCC 14400
 14401 ATCTGCTGTT AGAATAGCAA CAGTGTACTC ATGGGCTTAC GGTGATGATG ATAGCTCTTG GAACGAAGCC
 TGGTTGTTG CTAGGCAAAG GGCGCAATGTC 14500
 14501 AGCCGGAGG AGCTAAGGGT GATCACTCCC ATCTCAACTT CGACTAATT AGCGCATAGG TTGAGGGATC
 GTAGCACTCA AGTGAATAC TCAGGTACAT 14600
 14601 CCCTTGTCGG AGTGGCGAGG TATACCACAA TCTCCAACGA CAATCTCTCA TTTGTCATAT CAGATAAGAA
 GTTTGATACT AACTTTATAT ACCAACAAAGG 14700
 14701 AATGCTCTA GGGTTGGTG TTTAGAAC ATTGTTGCA CTCGAGAAAG ATACCGGATC ATCTAACACG
 GTATTACATC TTCACGTCGA AACAGATTGT 14800
 14801 TGCCTGATCC CGATGATAGA TCATCCCAGG ATACCCAGCT CCCGCAAGCT AGAGCTGAGG GCAGAGCTAT
 GTACCAACCC ATTGATATAT GATAATGCAC 14900
 14901 CTTTAATTGA CAGAGATGCA ACAAGGCTAT ACACCCAGAG CCATAGGAGG CACCTTGTGG AATTGTTAC
 ATGGTCCACA CCCCAACTAT ATCACATTTC 15000
 15001 AGCTAAGTC ACAGCACTAT CTATGATTGA CCTGGTAACA AAATTTGAGA AGGACCATAT GAATGAAATT
 TCAGCTCTCA TAGGGATGA CGATATCAAT 15100
 15101 AGTTTCATAA CTGAGTTCT GCTCATAGAG CCAAGATTAT TCACTATCTA CTTGGGCCAG TGTGCGGCCA
 TCAATTGGGC ATTGATGTA CATTATCATA 15200
 15201 GACCATCAGG GAAATATCAG ATGGGTGAGC TTGTTGTCATC GTTCCTTTCT AGAATGAGCA AAGGAGTGT
 TAAGGTGCTT GTCAATGCTC TAAGGCCACCC 15300
 15301 AAAGATCTAC AAGAAATTCT GGCATTGTGG TATTATAGAG CCTATCCATG GTCCTTCACT TGATGCTCAA
 AACATTGCA CAACTGTGTC CAACTGGT 15400
 15401 TACACATGCT ATATGACCTA CCTCGACCTG TTGTTGAATG AAGAGTTAGA AGAGTTCACA TTTCTCTTG
 GTGAAAGCGA CGAGGATGTA GTACCGGACA 15500
 15501 GATTGACAA CATCCAGCA AAACACTTAT GTGTTCTGGC AGATTGTAC TGTCAACCAG GGACCTGCC
 ACCAAATTGCA GGTCTAAGAC CGGTAGAGAA 15600
 15601 ATGTGCAGTT CTAACGGACC ATATCAAGGC AGAGGCTATG TTATCTCCAG CAGGATCTTC GTGGAACATA
 ATCCAATTA TTGAGGACCA TTACTCATGC 15700
 15701 TCTCTGACTT ATCTCCGGC AGGATCGATC AAACAGATAA GATTGAGAGT TGATCCAGGA TTCATTTTCG
 ACGCCCTCGC TGAGGTAAT GTCACTGCAG 15800
 15801 CAAAGATCGG CAGCAACAC ATCTCAAATA TGAGCATCAA GGCTTTCAGA CCCCCACACG ATGATGTTGC
 AAAATTGCTC AAAGATATCA ACACAAGCAA 15900
 15901 CGACACATCTT CCCATTTCAG GGGCAATCT CGCCAATTAT GAAATCCATG CTTTCCGCAG AATCGGGTTG
 AACATCATCG TTGCTACAA AGCTGTTGAG 16000
 16001 ATATCAACAT TAATTAGGAG ATGCCTTGAG CCAGGGGAGG ACGGCTTGTGTT CTTGGGTGAG GGATCGGGTT
 CTATGTTGAT CACTTATAAG GAGATACTTA 16100
 16101 AACTAAACAA GTGCTTCTAT AATAGTGGGG TTTCCGCCAA TTCTAGATCT GGTCAAAGGG AATTAGCAC
 CTATCCCTCC GAAGTGGG 16200
 16201 CAGAATGGGA GTAGGTAATA TTGCTAAAGT GCTCTTTAAC GGGAGGCCG AAGTCACGTG GGTAGGCAGT
 GTAGATTGCT TCAATTTCAT AGTTAGTAAT 16300
 16301 ATCCCTACCT CTAGTGTGGG GTTATCCAT TCAGATATAG AGACCTGCC TGACAAAGAT ACTATAGAGA
 AGCTAGAGGA ATTGGCAGCC ATCTTATCGA 16400
 16401 TGGCTCTGCT CCTGGCAAATAGGATCAA TACTGGTGTAT TAAGCTTATG CCTTCAGCG GGGATTTGT
 TCAGGGGATT TTAAGTTATG TAGGGTCTCA 16500
 16501 TTATAGAGAA GTGAACCTTG TATACCCCTAG ATACAGCAAC TTCACTATCTA CTGAATCTTA TTTGGTTATG
 ACAGATCTCA AGGCTAACCG GCTAATGAAAT 16600
 16601 CCTGAAAAGA TTAAGCAGCA GATAATTGAA TCATCTGTGA GGACTTCACC TGGACTTATA GGTCACATCC
 TATCCATTAA GCAACTAACG TGCTACAAAG 16700
 16701 CAATTGTGGG AGACGCAAGT AGTAGAGGTG ATATCAATCC TACTCTGAAA MAACCTACAC CTATAGAGCA
 GGTGCTGATC AATTGCGGGT TGGCAATTAA 16800
 16801 CGGACCTAAG CTGTGCAAAG AATTGATCCA CCATGATGTT GCCTCAGGGC AAGATGGATT GCTTAATTCT
 ATACTCATC TCTACAGGGGA GTTGGCAAGA 16900
 16901 TTCAAAGACA ACCAAAGAAG TCAACAAGGG ATGTTCCACG CTTACCCGT ATTGGTAAGT AGCAGGCAAC
 GAGAACTTAT ATCTAGGATC ACCCGAAAT 17000
 17001 TTTGGGGCA CATTCTCTT TACTCCGGGA ACAGAAAGTT GATAAATAAG TTTATCCAGA ATCTCAAGTC
 CGGCTATCTG ATACTAGACT TACACCAGAA 17100

Figure 24(Contd..)

17101	TATCTTCGTT	AAGAACATCAT	CCAAGTCAGA	GAAACAGATT	ATTATGACGG	GGGGTTTGAA	ACGTGAGTGG							
	GT	TTTAAGG	TAACAGTCAA	GGAGACAAA	17200									
17201	GAATGGTATA	AGTTAGTCGG	ATACAGTGCC	CTGATTAAGG	ACTAATTGGT	TGAACCTCGG	AACCTAAATC							
	CTGCCCTAGG	TGGTTAGGCA	TTATTTGCAA	17300										
17301	TATATTAAAG	AAAACATTGA	AAATAACGAAG	TTTCTATTCC	CAGCTTTGTC	TGGTggccgg	catggtccca							
	gcctccctcg	tggcgcggc	tgggcaacat	17400										
17401	tccgagggga	ccgtccccctc	ggtaatggcg	aatgggacGC	GGCCgatccg	gctgctaaca	aagcccggaaa							
	ggaagctgag	ttggctcgctg	ccacccgctga	17500										
17501	gcaataacta	gcataacccc	ttggggctc	taaacgggtc	ttgaggggtt	tttgctgaa	aggaggaact							
	atatccggat	GGGGCCGCAG	GTACCCCAGCT	17600										
17601	TTTGTTCCT	ttagtgggg	ttaattTCGA	GCTTGGCTA	ATCATGGTCA	TAGCTGTTTC	CTGTGTGAAA							
	TTGTTATCCG	CTCACAAATTC	CACACAACAT	17700										
17701	ACGAGCCGGA	AGCATAAAAGT	GTAAAACCTG	GGGTGCCTAA	TGAGTGAGCT	AACTCACATT	AATTGCGTTG							
	CGCTCACTGC	CGCGCTTCCA	GTCCGGAAAC	17800										
17801	CTGCTGTGCC	AGCTGCATTA	ATGAATCGGC	CAACGCGCGG	GGAGAGGCGG	TTTGCCTATT	GGGCCTCTT							
	CCGCTTCCTC	GCTCACTGAC	TCGCTGCGCT	17900										
17901	CGGTCGTTCC	GCTGCGGCGA	GCGGTATCG	CTCACTCAA	GGCGGTAATA	CGGTATCCA	CAGAACTCAGG							
	GGATAACGCA	GGAAAGAAC	TGTGAGCAA	18000										
18001	AGGCCAGCAA	AAGGCCAGGA	ACCGTAAAAA	GGCCGCGTTG	CTGGCGTTT	TCCATAGGCT	CCGCCCCCCT							
	GACGAGCAGTC	ACAAAAAATCG	ACGCTCAAGT	18100										
18101	CAGAGGTGGC	GAAACCCGAC	AGGACTATAA	AGATACCAGG	CGTTTCCCCC	TGGAAGCTCC	CTCGTGCCT							
	CTCCCTGTTCC	GACCTGCCCC	CTTACCGGAT	18200										
18201	ACCTGTCCGC	CTTTCTCCCT	TCGGGAAGCG	TGGCGCTTTC	TCATAGCTCA	CGCTGTAGGT	ATCTCAGTTC							
	GGTGAGGTC	GTTCGCTCCA	AGCTGGGCTG	18300										
18301	TGTGCACGCA	CCCCCGCTTC	AGCCCGACCG	CTGCGCCTTA	TCCGGTAACT	ATCGTCTTGA	GTCCAACCCG							
	GTAAGACACG	ACTTATCGCC	ACTGGCAGCA	18400										
18401	GCCACTGGTA	ACAGGATTAG	CAGAGCGAGG	TATGTAGGCG	GTGCTACAGA	GTTCTTGAAG	TGGTGGCCTA							
	ACTACGGCTA	CACTAGAAGG	ACAGTATTTG	18500										
18501	GTATCTGCGC	TCTGCTGAAG	CCAGTTACCT	TCGGAAAAAG	AGTTGGTAGC	TCTTGATCCG	GCAAACAAAC							
	CACCGCTGGT	ACGGGTGGTT	TTTTGTGTTG	18600										
18601	CHAGCAGCAG	ATTACGGCA	GAAAAAAAGG	ATCTCAAGAA	GATCCTTGA	TCTTTCTAC	GGGGTCTGAC							
	GCTCAGTGGA	ACGAAAACCTC	ACGTTAAGGG	18700										
18701	ATTTGGTCA	TGAGATTATC	AAAAAGGATC	TTCACCTAGA	TCCTTTAAA	TTAAAAATGA	AGTTTTAAAT							
	CAATCTAAAG	TATATATGAG	TAACACTGGT	18800										
18801	CTGACAGTTA	CCAATGCTTA	ATCAGTGAGG	CACCTATCTC	AGCGATCTGT	CTATTTCGTT	CATCCATAGT							
	TGCCCTGACTC	CCCGCTGTGT	AGATAACTAC	18900										
18901	GATACGGGAG	GGCTTACCAT	CTGGCCCCAG	TGCTGCAATG	ATACCGCGAG	ACCCACGCTC	ACCGGCTCCA							
	GATTTATCAG	CAATAAACCA	GCCAGCCGGA	19000										
19001	AGGGCCGAGC	GCAGAACTGG	TCCTGCAACT	TTATCCGCCT	CCATCCAGTC	TATTAATTGT	TGCCGGGAAG							
	CTAGAGTAAG	TAGTCGCCA	GTTAATAGTT	19100										
19101	TGCGAACGT	TGTTGCCATT	GCTACAGGCA	TCGTGGTGTC	ACGCTCGTC	TTTGGTATGG	CTTCATTCA							
	CTCCGGTTCC	CAACGATCAA	GGCGAGTTAC	19200										
19201	ATGATCCCCC	ATGTTGTGCA	AAAAAGCGGT	TAGCTCCTTC	GGTCCTCCGA	TCGTTGTCAG	AAGTAAGTTG							
	GCCGCAGTGT	TATCCTCAT	GGTTATGGCA	19300										
19301	GCACTGCATA	ATTCTCTTAC	TGTCTATGCCA	TCCGTAAGAT	GCTTTCTGT	GACTGGTGAG	TACTCAACCA							
	AGTCATTCTG	AGAATAGTGT	ATGCGGCAC	19400										
19401	CGAGTTGCTC	TTGCCGGCG	TCAATACGGG	ATAATACCGC	GCCACATAGC	AGAACTTAA	AAGTGCCTAT							
	CATTGGAAAA	CGTTCTCGG	GGCGAAAAC	19500										
19501	CTCAAGGATC	TTACCGCTGT	TGAGATCCAG	TTCGATGTAA	CCCACTCGTG	CACCCAACTG	ATCTTCAGCA							
	TCTTTTACTT	TCACACCGGT	TTCTGGGTGA	19600										
19601	GCAAAACACG	GAAGGCAAAA	TGCCGCAAAA	AAGGGAATAA	GGGCGACACG	GAAATGTTGA	ATACTCATAAC							
	TCTTCCTTT	TCAATATTAT	TGAAGCATT	19700										
19701	ATCAGGGTTA	TTGTCTCATG	AGCGGATACA	TATTTGAATG	TATTTAGAAA	AATAAACAAA	TAGGGTTCC							
	GCGCACATT	CCCCGAAAAG	TGC	19793										
		10		20		30		40		50		60		70
		80		90		100								

Figure 24 (Contd..)

	80	10	20	30	40	50	60	70
1	CACCTAAATT	GTAAGCGTTA	ATATTTGTT	AAAATTCGGC	TTAAATTTT	GTTAAATCAG	CTCATTTTT	
AACCAATAGG	CCGAAATCGG	CAAATCCCT	100					
101	TATAAATCAA	AAGAAATPAGAC	CGAGATAGGG	TTGAGTGTG	TTCCAGTTG	GAACAAGAGT	CCACTATTAA	
AGAACGCTGA	CTCCAACGTC	AAAGGGCGAA	200					
201	AAACCGTCTA	TCAGGGCGAT	GGCCCACATC	GTGAACCATC	ACCCCTAAC	AGTTTTTGG	GGTCGAGGTG	
CCGTAAGCA	CTAAATCGGA	ACCCCTAAAGG	300					
301	GAGCCCCCGA	TTTAGAGCTT	GACGGGGAAA	GCCggccatt	taggccaTAG	GGCGCTGGCA	AGTGTAGCGG	
TCACGCTGCG	CGTAACACC	ACACCCGCGG	400					
401	CGCTTAAATGC	GCCGCTAACAG	GGCCTGCTCC	ATTGCCATT	CAGGCTGCGC	AACTGTTGGG	AAGGGCGATC	
GGTGCAGGCC	TCTTCGCTAT	TACGCCAGCT	500					
501	GCGAAAGGG	GGATGTGCTG	CAAGGCGATT	AAGTTGGGT	ACGCCAGGGT	TTTCCCAGTC	ACGACGTTGT	
AAAACGACGG	CCAGTGAATT	Gtaatacgcac	600					
601	tcaactataAC	CAAACAAAGT	TGGGTAAGGA	TAGTTCAATC	AATGATCATC	TTCTAGTGCA	CTTAGGATTTC	
AAGATCCTAT	TATCAGGGAC	AAGAGCAGGA	700					
701	TTAGGGATAT	CTGAGATGGC	CACACTTTA	AGGAGCTTAG	CATTGTTCAA	AAGAAACAAG	GACAAACAC	
CCATTACATC	AGGATCCGGT	GGAGCCATCA	800					
801	GAGGAATCAA	ACACATTATT	ATAGTACCAA	TCCCTGGAGA	TTCCCTCAATT	ACCACTCGAT	CCAGACTTCT	
GGACCCGTTG	GTCAAGTTAA	TTGGAAACCC	900					
901	GGATGTGAGC	GGGCCAAC	TAACAGGGC	ACTAATAGGT	ATATTATCCT	TATTTGTGGA	GTCTCCAGGT	
CAATTGATTC	AGAGGATCAC	CGATGACCT	1000					
1001	GACGTTAGCA	TAAGGCTGTT	AGAGGTTGTC	CAGAGTGACC	AGTCACAATC	TGGCCTTACC	TTCGCATCAA	
GAGGTACCAA	CATGGAGGAT	GAGGCGGACC	1100					
1101	AATACTTTTC	ACATGATGAT	CCAATTAGTA	GTGATCAATC	CAGGTTCGGA	TGGITCGAGA	ACAAGGAAAT	
CTCAGATATT	GAAGTGAAG	ACCCCTGAGGG	1200					
1201	ATTCAACATC	ATTCTGGTA	CCATCCTAAC	CCAAATTG	GTCTTCTCG	CAAAGCGGT	TACGGCCCCA	
GACACGGCAG	CTGATTCCGA	GCTAAGAAGG	1300					
1301	TGGATAAAGT	ACACCCAA	AAGAAGGGTA	GTTGGTGAAT	TTAGATTGGA	GAGAAAATGG	TTGGATGTGG	
TGAGGAACAG	GATTGCCAG	GACCTCTCT	1400					
1401	TACGCCGATT	CATGGTCGCT	CTAACCTCTG	ATATCAAGAG	AACACCCGGA	AACAAACCA	GGATTGCTGA	
AATGATATGT	GACATTGATA	CATATATCGT	1500					
1501	AGAGGCAGGA	TTAGCCAGT	TTATCCTGAC	TATTAAGTTT	GGGATAGAAA	CTATGTATCC	TGCTCTTGG	
CTGCATGAAT	TTGCTGGTGA	GTATCCACA	1600					
1601	CTTGAGTCCT	TGATGAACCT	TTACCGACAA	ATGGGGAAA	CTGCACCCCTA	CATGGTAATC	CTGGAGAACT	
CAATTCAAGAA	CAAGTTCAGT	GCAGGATCAT	1700					
1701	ACCCCTGCT	CTGGAGCTAT	GCCATGGAG	TAGGAGTGG	ACTTGAAAAC	TCCATGGGG	GTGAACTCGGT	
TGGCCGATCT	TACTTGTGATC	CAGCATATT	1800					
1801	TAGATTAGGG	CAAGAGATGG	TAAGGAGGTC	AGCTGGAAAG	GTCAGTTCCA	CATGGCATC	TGAACCTCGGT	
ATCACTGCCG	AGGATGCAAG	GCTTGTTC	1900					
1901	GAGATTGCAA	TGCACTACTAC	TGAGGACAAG	ATCACTGAG	CGGTTGGACC	CAGACAAGCC	CAAGTATCAT	
TTCTACACGG	TGATCAAAGT	GAGAATGAGC	2000					
2001	TACCGAGATT	GGGGGGCAAG	GAAGATAGGA	GGGTCAAAC	GAGTCGAGGA	GAAGCCAGGG	AGAGCTACAG	
AGAAACCGGG	CCCAGCAGAG	CAAGTGTG	2100					
2101	GAGAGCTGCC	CATCTTCAA	CCGGCACACC	CCTAGACATT	GACACTGCAT	CGGAGTCCAG	CCAAGATCCG	
CAGGACACTG	GAAGGTCAGC	TGACGGCTCTG	2200					
2201	CTTAGGCTGC	AAGCCATGGC	AGGAATCTG	GAAGAACAAAG	GCTCAGACAC	GGACACCCCT	ATAGTGTACA	
ATGACAGAAA	TCTTCTAGAC	TAGGTGCGAG	2300					
2301	AGGCCGAGGG	CCAGAACAC	ATCCGCCTAC	CCTCCATCAT	TGTTATAAAA	AACTTAGGAA	CCAGGTCCAC	
ACAGCCGCA	GCCCATCAAC	CATCCACTCC	2400					
2401	CACGATTGGA	GCCAATGTTA	GAAGAGCAGG	CACGCCATGT	CAAAAACGGA	CTGGAATGCA	TCCGGCTCT	
CAAGGCCGAG	CCCATCGGCT	CACTGGCCAT	2500					
2501	CGAGGAAGCT	ATGGCAGCAT	GGTCAGAAAT	ATCAGACAAC	CCAGGACAGG	AGCGAGCCAC	CTGCAGGGAA	
GAGAAGGCGAG	GCAGTTCGGG	TCTCAGAAA	2600					
2601	CCATGCTCT	CAGCAATTGG	ATCAACTGAA	GGCGGTGCAC	CTCGCATCCG	CGGTCAAGGA	CCTGGAGAGA	
GCGATGACGA	CGCTGAAACT	TTGGGAATCC	2700					
2701	CCCCAAGAAA	TCTCCAGCA	TCAAGCAGTC	GGTTACAGTG	TTATTACGTT	TATGATCACA	GGGGTGAAGC	
GGTTAAGGGA	ATCCAAGATG	CTGACTCTAT	2800					
2801	CATGGTTCAA	TCAGGCCCTG	ATGGTGTAG	CACCCCTCTCA	GGAGGAGACA	ATGAATCTGA	AAACAGCGAT	
GTGGATATTG	GCAGAACCTGA	TACCGAGGG	2900					
2901	TATGCTATCA	CTGACCGGGG	ATCTGCTCCC	ATCTCTATGG	GGTTCAAGGC	TTCTGATGTT	GAAGACTGCAG	
AAGGAGGGGA	GATCCACCG	CTCCCTGAGAC	3000					
3001	TCCAATCCAG	AGGCAACAC	TTTCCGAAGC	TTGGGAAAAC	TCTCAATGTT	CCTCCGCCCC	CGGACCCCGG	
TAGGGCCAGC	ACTTCCGGGA	CACCCATTAA	3100					
3101	AAAGGGCACA	GACCGGAGAT	TAGCCTCATT	TGGAACGGAG	ATCGCGTCTT	TATTGACAGG	TGGTGCAACC	
CAATGTGCTC	GAAAGTCACC	CTCGGAACCA	3200					
3201	TCAGGGCCAG	GTGCACCTGC	GGGGAAATGTC	CCCGAGTGTG	TGAGCAATGC	CGCACTGATA	CAGGAGTGGA	
CACCGAATC	TGGTACCCACA	ATCTCCCGA	3300					
3301	GATCCCAGAA	TAATGAAGAA	GGGGGAGACT	ATTATGATGA	TGAGCTGTC	TCTGATGTC	AAGATATTAA	

Figure 25

AACAGCCTTG GCCAAATAC ACGAGGATAA 3400
 3401 TCAGAAGATA ATCTCCAAGC TAGAATCACT GCTGTTATTG AAGGGAGAAG TTGAGTCAAT TAAGAACGAG
 ATCAACAGGC AAAATATCAC CATATCCACC 3500
 3501 CTGGAAGGAC ACCTCTCAAG CATCATGATC GCCATTCTG GACTTGGGAA GGATCCCAAC GACCCCAGT
 CAGATGCGA AATCAATCCC GACTTGAAAC 3600
 3601 CCATCATAGG CAGAGATCA GGCGGAGCAC TGGCCGAAGT TCTCAAGAAA CCCGTTGCCA GCCGACA
 CCAAGGAATG ACAAATGGAC GGACCCAGTTC 3700
 3701 CAGAGGACAG CTGCTGAAGG AATTTCAGCT AAAGCCGATC GGGAAAAAGA TGAGCTCAGC CGTCGGGTTT
 GTTCCCTGACA CGGGCCCTGC ATCACGCACT 3800
 3801 GTAATCCGCT CCATTATAAA ATCCAGCCG CTAGAGGAGG ATCGGAAGCG TTACCTGATG ACTCTCCTTG
 ATGATATCAA AGGAGCCAAT GATCTTGCA 3900
 3901 AGTCCACCA GATGCTGATG AAGATAATAA TGAAGTAGCT ACAGCTCAAC TTACCTGCCA ACCCCATGCC
 AGTCGACCA actagtACAA CCTAAATCCA 4000
 4001 TCATAAAAAA CTTAGGAGCA AAGTGATTGC CTCCCAAGTT CCACAATGAC AGAGATCTAC GACTTCGACA
 AGTCGGCATG GGACATCAA GGGTCGATCG 4100
 4101 CTCCGATACA ACCCACCAAC TACAGTGTG GCAGGCTGGT GCCCCAGGTC AGAGTCATAG ATCCTGGTCT
 AGGCAGCAGG AAGGATGAAT GCTTTATGTA 4200
 4201 CATGTTCTG CTGGGGTTG TTGAGGACAG GGATTCCCTA GGGCCTCCAA TCGGGCGAGC ATTTGGTCC
 CTGCCCCCTAG GTGTTGGCAG ATCCACAGCA 4300
 4301 AAGCCCGAAA AACTCTCAA AGAGGCCACT GAGCTTGACA TAGTTGTTAG ACGTACAGCA GGGCTCAATG
 AAAAATGGT GTTCTACAAC AACACCCAC 4400
 4401 TAATCTCCT CACACCTGG AGAAAGTCC TAACAACAGG GAGTGTCTC AACGCAAACC AAGTGTGCAA
 TGCAGGTTAAT CTGATACCGC TCGATACCCC 4500
 4501 CGAGGGTTC CGTGTGTTT ATATGAGCAT CACCCGTCTT TCGGATAACG GGTATTACAC CGTTCTAGA
 AGAATGCTGG AATTCAAGTC GGTCAATGCA 4600
 4601 GTGGCCTTC ACCTGCTGGT GACCCCTTAGG ATTGACAAGG CGATAGGCC TGGGAAGATC ATCGACAATA
 CAGAGCAACT CCTGAGGCA ACATTATAG 4700
 4701 TCCACATCGG GAACCTCAGG AGAAAGAAGA GTGAAGTCTA CTCTGCCGAT TATTGAAAAA TGAAAATCGA
 AAAGATGGGC CTGGTTTTG CACTTGGTGG 4800
 4801 GATAGGGGGC ACCAGTCTTC ACATTTAGAG CACAGGCAAA ATGAGCAAGA CTCTCAATGC ACAACTCGGG
 TTCAAGAAGA CCTTATGTTA CCCGCTGATG 4900
 4901 GATATCAATG AAGACCTTAA TCGATTACTC TGGAGGAGCA GATGCAAGAT AGTAAGAAC CAGGAGTTT
 TGCAGCCATC AGTCCCTCAA GAATTCCGCA 5000
 5001 TTTACGACGA CGTGATCATA AATGATGACC AAGGACTATT CAAAGTTCTG TAGACCGTAG TGCCCAGCAA
 TGCCCCGAAA CGACCCCTC CACAATGACA 5100
 5101 GCGCAAGGG CGCGACACCC AAGCCCTC CGAAAGACTC CACGGACCAA GCGAGAGGCC AGCCAGCAGC
 CGACGGCA CGCGAACACC AGGGGGCCCC 5200
 5201 AGCACAGAAC AGCCCTGACA CAAGGCCACC ACCAGCCACC CCAATCTGCA TCCTCTCGT GGGACCCCCG
 AGGACCAACC CCCAAGGCTG CCCCCGATCC 5300
 5301 AAACCAACCA CGCGATCCCC ACCACCCCCG GGAAAGAAAC CCCCAGCAAT TGGAAAGGCC CTCCCCCTCT
 TCCTCAACAC AAGAACCTCA CAACCGAAC 5400
 5401 GCACAAGCGA CCGAGGTGAC CCAACCGCAG GCATCCGACT CCCTAGACAG ATCCTCTCTC CCCGCAAAC
 TAAACAAAC TTAGGGCCAA GGACATACATA 5500
 5501 CACCCAACAG AACCCAGACC CGGGCCCGAC GCGCCGCGCC CCCAACCCCC GACAACCAGA GGGAGCCCC
 AACCAATCCC GCGGCTCCC CGGTGCCA 5600
 5601 CAGGCAAGGA CACCAACCCC CGAACAGACC CAGCACCAA CCATGACAA TCCAAGACGG GGGGGCCCC
 CCAAAAAAAA GCCCCCAGGG GCGCACAGCC 5700
 5701 AGCACCGCGA GGAAGCCAC CCACCCCCACA CACGACCACG GCAACCAAAC CAGAACCCAG ACCACCCCTGG
 GCCACCCAGCT CCCAGACTCG GCGATCACCC 5800
 5801 CGCAGAAAGG AAAGGCCACA ACCCGCGCAC CCCAGCCCCG ATCCGGGG GAGGCCACCA ACCCGAACCA
 GCACCCAAAGA CGGATCCCCG AAGGACCCCC 5900
 5901 GAACCGAAA GGACATCAGT ATCCACACAGC CTCTCCAAGT CCCCCGGTCT CCTCCTCTC TCGAAGGGAC
 CAAAGATCA ATCCACCCACA CCCGACGACA 6000
 6001 CTCAACTCCC CACCCCTAAA GGAGACACCG GGAATCCCAG AATCAAGACT CATCCAATGT CCATCATGGG
 TCTCAAGGT AACGTCCTG CCATCATCAT 6100
 6101 GGCAGTACTG TTAACTCTCC AAACACCCAC CGGTCAAATC CATTGGGCC ATCTCTCTAA GATAGGGGTG
 GTAGGAATAG GAAGTGCAG CTACAAAGTT 6200
 6201 ATGACTCGTT CCAGCCATCA ATCATTAGTC ATAAAATTAA TGCCCAATAT AACTCTCTC AATAACTGCA
 CGAGGGTAGA GATTGCGAGA TACAGGAGAC 6300
 6301 TACTGAGAAC AGTTTGGAA CCAATTAGAG ATGCACTTAA TGCAATGACC CAGAAATATAA GACCGGTTCA
 GAGTGTAGCT TCAAGTAGGA GACAAAGAG 6400
 6401 ATTTGCGGGA GTAGTCTGG CAGGTGCGGC CCTAGGCCTT GCCACAGCTG CTCAGATAAC GGCGGGCATT
 GCACCTCACC AGTCCATGCT GAACTCTCAA 6500
 6501 GCCATCGACA ATCTGAGAGC GAGCCTGGAA ACTACTAATC AGGCAATTGA GGCAATCAGA CAAGCAGGGC
 AGGAGATGAT ATTGGCTGTT CAGGGTGTCC 6600
 6601 AAGACTACAT CAATAATGAG CTGATACCGT CTATGAACCA ACTATCTTGT GATTTAATCG GCCAGAAGCT
 CGGGCTCAAAT TTGCTCAGAT ACTATACAGA 6700
 6701 AACCTGTCA TTATTTGGCC CGAGTTACG GGACCCATA TCTGCGGAGA TATCTATCCA GGCTTGAGC

Figure 25(Contd..)

TATGCGCTTG GAGGAGACAT CAATAAGGTG 6800
 6801 TTAGAAAAGC TCGGATACAG TGGAGGTGAT TTACTGGCA TCTTAGAGAG CAGAGGAATA AAGGCCCGGA
 TAACTCACGT CGACACAGAG TCCCTACTTCA 6900
 6901 TTGTCCTCG TATAGCTAT CCCGACCGTGT CCGAGATTAA GGGGGTGATT GTCCACCGGC TAGAGGGGT
 CTCGTACAC ATAGGCCTC AAAGAGTGGTA 7000
 7001 TACCACTGTG CCCAAGTATG TTGCAACCCA AGGGTACCTT ATCTCGAATT TTGATGAGTC ATCGTGTACT
 TTCATGCCAG AGGGGACTGT GTGAGCCAA 7100
 7101 ATATGCCCTGT ACCCGATGAG TCCTCTGCTC CAAGAATGCC TCCGGGGTA CACCAAGTCC TGTGCTCGTA
 CACTCGTATC CGGGTCTTTT GGGAACCGGT 7200
 7201 TCATTTTAC ACAAGGAAAC CTAATAGCCA ATTGTGCATC AATCCTTGC AAGTGTACA CAACAGGAAC
 GATCATTAAT CAAGACCCCTG ACAAGATCCT 7300
 7301 AACATACATT GCTGCCGATC ACTGCCCGT AGTCGAGGT AACGGCGTGA CCATCCAAGT CGGGGAGCAGG
 AGGTATCCAG ACGCTGTGTA CTTGACAGA 7400
 7401 ATTGACCTCG GTCCTCCAT ATCATGGAG AGGTTGGACG TAGGGACAAA TCTGGGAAT GCAATTGCTA
 AGTTGGAGGA TGCCAAGGA TTGTTGGAGT 7500
 7501 CATCGGACCA GATATTGAGG AGTATGAAAG GTTATCGAG CACTAGCATA GTCTACATCC TGATTGCAGT
 GTGTCTTGGA GGGTTGATAG GGATCCCCGC 7600
 7601 TTTAATATGT TGCTGCAGGG GGCCTGTAA CAAAAAGGA GAACAAGTTG GTATGTCAAG ACCAGGCCTA
 AAGCCTGATC TTACGGGAAC ATCAAATCC 7700
 7701 TATGTAAGGT CGCTCTGATC CTCTACAACT CTTGAAACAC AAATGTCCC CAAGTCTCCT CTTCGTCATC
 AAGCAACCAC CGCACCCAGC ATCAAGGCCA 7800
 7801 CCTGAAATTA TCTCCGGCTT CCCTCTGGCC GAACAATATC GGTAGTTAAT TAAAACCTAG GGTGCAAGAT
 CATCCACAAAT GTCACACAA CGAGACCGGA 7900
 7901 TAAATGCCTT CTACAAAGAT AACCCCCATC CCAAGGGAAG TAGGATAGTC ATTAACAGAG AACATCTTAT
 GATTGATAGA CCTTTATGTT TGCTGGTGT 8000
 8001 TCTGTTTGTGATGTTGATCGG GTTGCTAGCC ATTGCGAGCA TTAGACTTCA TCGGGCAGCC
 ATCTACACCC CAGAGATCCA TAAAGGCCTC 8100
 8101 AGCACCAATC TAGATGTAAC TAATCAATC GAGCATCAGG TCAAGGACGT GCTGACACCA CTCTTCAAA
 TCATCGGTGA TGAAGTGGGC CTGAGGACAC 8200
 8201 CTCAGAGATT CACTGACCTA GTGAAATTCA TCTCTGACAA GATTAAATTC CTTAATCCGG ATAGGGAGTA
 CGACTTCAGA GATCTCACTT GGTTATCAC 8300
 8301 CCCGCCAGAG AGAATCAAAT TGGATTATGA TCAATACTGT GCAGATGTGG CTGCTGAAGA CCTCATGAAT
 GCATTGGTGA ACTCAACTCT ACTGGAGACAG 8400
 8401 AGAACACCA ATCAAGTCTCCT AGCTGTCTCA AAGGGAAACT GCTCAGGGCC CACTACAATC AGAGGTCAAT
 TCTCAACAT GTCGCTGTCC CTGTTAGACT 8500
 8501 TGTATTTAGG TCGAGGTTAC AATGTGTCACT CTATAGTCAC TATGACATCC CAGGGATGT ATGGGGAAAC
 TTACCTAGTC GAAAAGCTA ATCTGAGCAG 8600
 8601 CAAAAGGTCA GAGTTGTCAC AACTGAGCAT GTACCGAGTG TTTGAAGTAG GTGTTATCAG AAATCCGGT
 TTGGGGGCTC CGGTGTTCCA TATGACAAAC 8700
 8701 TATCTTGAGC AACCAGTCAG TAATGATCTC AGCAACTGTA TGGTGGCTTT GGGGGAGTC AAACTCGCAG
 CCCTTTGTCA CGGGGAAGAT TCTATCACAA 8800
 8801 TTCCCTATCA GGGATCAGGG AAAGGTGTC GCTTCCAGCT CGTCAAGCTA GGTGTCTGGA ATCCCCAAC
 CGACATGCAA TCCTGGGTCC CCTTATCAC 8900
 8901 GGATGATCCA GTGATAGACA GGCTTACCT CTCATCTCAC AGAGGTGTTA TCGCTGACAA TCAAGCAAA
 TGGGCTGTCC CGAACACACG AACAGATGAC 9000
 9001 AAGTTGCGAA TGGAGACATG CTCCAACAG GCGTGTAAAG GTAAAATCCA AGCACTCTGC GAGAATCCCG
 ACTGGGCACC ATTGAAGGAT AACAGGATTC 9100
 9101 CTTCATACGG GGTCTTGTC GTTGATCTGA GTCTGACAGT TGAGCTTAA ATCAAAATTG CTTGGGATT
 CGGGCCATTG ATCACACACG GTTCAGGGAT 9200
 9201 GGACCTATAC AAATCCAAC ACACAAATGT GTATTGGCTG ACTATCCGC CAATGAAGAA CCTAGCCTTA
 GTGTAAATCA ACACATGGA GTGATACCG 9300
 9301 AGATTCAGG TTAGTCCTA CCTCTTCACT GTCCCAATTAA AGGAAGCAGG CGAAGACTGC CATGCCCAA
 CATACTTACCG TCGGGAGGTG GATGGTGTG 9400
 9401 TCAAACCTAG TTCCAATCTG GTGATCTAC CTGGTCAAGA TCTCCAATAT GTTTGGCAA CCTACGATAC
 TTCCAGGGTT GAACATGCTG TGTTTATTA 9500
 9501 CGTTTACAGC CCAGGCCGCT CATTTCCTA CTTTTATCCT TTTAGGTGTC CTATAAAGGG GGTCCCCATC
 GAATTACAGG TGGAAATGCTT CACATGGAC 9600
 9601 CAAAACCTCT GGTGCCGTC CTTCTGTGTG CTTGCCGACT CAGAACTGG TGGACATATC ACTCACTCTG
 GGATGGTGGG CATGGGAGTC AGCTGCACAG 9700
 9701 TCACCCGGGA AGATGAAACC AATGCCAGAT AGGGCTGCTA GTGAACCAAT CACATGATGT CACCCAGACA
 TCAGGCATAC CCActagtccatc 9800
 9801 attgttataa aaaacttagg aaccaggatcc acacagccgc cagcccatca AcgcgtATCT TCACCGGTGA
 TCTATCGCGT acgtacgcgc catgataaa 9900
 9901 ggagaagaac ttttcaactgg agttgtccca attcttggat aatttagatgg tgatgttaat gggcacaaat
 tttctgtcag tggaggggt gaaggtatg 10000
 10001 caacatacg aaaaactacc cttaaatcta ttgcactac tggaaaacta cctgttccat ggccaacact
 tgtcaactt ttcaactatg gtgtcaatg 10100
 10101 cttttcaaga tacccagatc atatgaaacg gcatgacttt ttcaagagtg ccatgcccga aggttacgt
 cagggaaatgaa ctatattttt caaagatgac 10200

Figure 25(Contd..)

10201 ggaaactaca agacacgtgc tgaagtcaag tttgaaggta atacccttgt taatagaatc gagttaaaag
 gtattgatt taaagaagat ggaacatcc 10300
 10301 ttggacacaa attggataac aactataact cacacaatgt atacatcatg gcagacaaac aaaagaatgg
 aatcagatc aacttacaaa tttagacacaa 10400
 10401 cattgaagat ggaagcggtc aactagcaga ccattatcaa caaaaacttc caattggcga tggccctgtc
 cttttaccag acaaccatca cctgtccaca 10500
 10501 caatctgccc ttgcgaaaa tcccaacgaa aagagagacc acatggtcct tcttgagtt gtaacagctg
 ctgggattac acatggcatg gatgaaat 10600
 10601 acaaataatgc agcgcgcgc gctgacgtct cgcgatgata ctatgtgaa ATAGACATCA GAATTAAGAA
 AAACGTAGGG TCCAAGTGGT TCCCCGTTAT 10700
 10701 GGACTCGCTA TCTGTCAACC AGATCTTATA CCCTGAAGTT CACCTAGATA GCCCCGATAGT TACCAATAAG
 ATAGTAGCCA TCCTGGAGTA TGCTCGAGTC 10800
 10801 CCTCACCGCTT ACAGCCCTGGA GGACCCCTACA CTGTGTCAGA ACATCAAGCA CCCGCTAAAA AACGGATTTT
 CCAACCAAAT GATTATAAAC AATGTGGAAG 10900
 10901 TTGGGAATGT CATCAAGTCC AAGCTTAGGA GTTATCCGGC CCACTCTCAT ATTCCATATC CAAATTGTAA
 TCAGGATTTA TTTAACATAG AAGACAAAAGA 11000
 11001 GTCAACGAGG AAGATCCGTG AACTCCTCAA AAAGGGGAAT TCGCTGTACT CCAAAGTCAG TGATAAGGTT
 TTCCAATCCT TAAGGGACAC TAATCACCG 11100
 11101 CTTGGCCTAG GCTCCGAATT GAGGGAGGAC ATCAAGGAGA AAGTTATTAA CTTGGGAGTT TACATGCACA
 GCTCCCACTG GTTTGAGCCC TTTCTGTTT 11200
 11201 GGTTTACAGT CAAGACTGAG ATGAGGTCA TGATTAAATC ACAAAACCAT ACTTGCCATA GGAGGAGACA
 CACACCTGTA TTCTTCACTG GTAGTTCACT 11300
 11301 TGAGTTGCTA ATCTCTCGT ACCTTGTTC TATAATCAGT AAAGAGTCTC AACATGTATA TTACCTGACA
 TTGAACTGG TTTTGATGTA TTGATGATGTC 11400
 11401 ATAGAGGGGA GGTAAATGAC AGAGACCCGT ATGACTATTG ATGCTAGGTA TACAGAGCTT CTAGGAAGAG
 TCAGATACAT GTGGAAACTG ATAGATGGTT 11500
 11501 TCTTCCTGCACTCGGAAAT CCAACTTATC AAATTGTAGC AATGCTGGAG CCTCTTCAC TTGCTTACCT
 GCAGCTGAGG GATATAACAG TAGAAACTCAG 11600
 11601 AGGTGCTTTC CTTAACCACT GCTTACTGTA AATAATGAT GTTCTTGACC AAAACGGGTT TTCTGATGAA
 GCTACTTATC ATGACTTAAT TGAACTCTA 11700
 11701 GATTACACT TGATACTGAG TGACACATAC CTGACAGGGG AGATTTCTC ATTTTCAGA AGTTTCCGCC
 ACCCAGACT TGAAAGCAGTA ACGGCTGCTG 11800
 11801 AAAATGTTAG GAAATACATG AATCAGCCTA AAGTCATTGT GTATGAGACT CTGATGAAAG GTCATGCCAT
 ATTTTGTTGA ATCATAATCA ACGGCTATCG 11900
 11901 TGACAGGCAC GGAGGCAGTT GGCCACCGCT GACCCCTCCC CTGCATGCTG CAGACACAAT CCGGAATGCT
 CAAGCTTCAG GTGAAAGGTT AACACATGAG 12000
 12001 CAGTGCCTTG ATAACCTGAA ATCTTTGCT GGAGTGAAT TTGGCTGCTT TATGCCTTT AGCCTGGATA
 GTGATCTGAC AATGTACCTA AAGGCAAGG 12100
 12101 CACTTGCTGC TCTCCAAAGG GAATGGGATT CAGTTTACCC GAAAGAGTTC CTGCGTTACG ACCCTCCAA
 GGGAAACGGG TCACGGAGGC TTGTAGATGT 12200
 12201 TTTCCTTAAT GATTGAGCT TTGACCCATA TGATGTGATA ATGTATCTG TAAGTGGAGC TTACCTCCAT
 GACCCCTGAGT TCAACCTGTC TTACAGCCTG 12300
 12301 AAAGAAAAAGG AGATCAAGGA AACAGGTAGA CTTTTGCTA AAATGACTTA CAAATGAGG GCATGCCAAG
 TGATTGCTGA AAATCTAATC TCAACGGGA 12400
 12401 TTGCAAATA TTTAAAGGAC AATGGGATGG CCAAGGATGA GCACGATTG ACTAAGGCAC TCCACACTCT
 AGCTGTCTCA GGAGTCCCCA AAGATCTAA 12500
 12501 AGAAAGTCAC AGGGGGGGC CAGTCTTAAA AACCTACTCC CGAAGCCAG TCCACACAAG TACCAAGGAAC
 GTGAGAGCAG CAAAAAGGTT TATAGGGTTC 12600
 12601 CCTCAAGTAA TTCGGCAGGA CCAAGACACT GATCATCCGG AGAATATGGA AGCTTACGAG ACAGTCAGTG
 CATTATACAC GACTGATCTC AAGAAGTACT 12700
 12701 GCCTTAATTG GAGATATGAG ACCATCAGCT TGTGTCACA GAGGCTAAAT GAGATTACG GATTGCCCTC
 ATTTTCCAG TGGCTGCATA AGAGGCTTGA 12800
 12801 GACCTCTGTC CTGTATGTA GTGACCCCTCA TTGCCCCCCC GACCTTGACG CCCATATCCC GTTATATAAA
 GTCCCCAATG ATCAAACTCTT CATTAAAGTAC 12900
 12901 CCTATGGGAG GTATAGAAGG GTATGTCAAG AAGCTGTGGA CCATCAGCAC CATTCCCTAT CTATACCTGG
 CTGCTTATGAG GAGGGAGGAGA AGGGCTTGA 13000
 13001 CGTTAGTGCAG AGGGGACAAAT CAGACCATAG CCGTAACAAA AAGGGTACCC AGCACATGGC CCTACAACCT
 TAAGAAACGG GAAGCTGCTA GAGTAACACTAG 13100
 13101 AGATTACTTT GTAATTCTTA GGCAAGGCT ACATGATATT GGCCATCACC TCAAGGAAA TGAGACAATT
 GTTTCATCAC ATTTTTTGT CTATCAAAA 13200
 13201 GGAATATATTG ATGATGGCT ACTTGTGTC CAATCACTCA AGAGCATCGC AAGATGTGTA TTCTGGTCAG
 AGACTATAGT TGATGAAACA AGGGCAGCAT 13300
 13301 GCACTAAATAT TGCTACAAAC ATGGCTAAAA GCATCGAGAG AGGTTATGAC CGTTACCTTG CATATTCCCT
 GAACGTCTTA AAAGTGATAC AGCAAATTCT 13400
 13401 GATCTCTCTT GGCTTCACAA TCAATTCAAC CATGACCCGG GATGTAGTCA TACCCCTCCT CACAAACAAAC
 GACCTCTTAA TAAGGATGGC ACTCTTGCCCC 13500
 13501 GCTCCTATTG GGGGGATGAA TTATCTGAAT ATGAGCAGGC TGTTGTCA GAAACATCGGT GATCCAGTAA
 CATCATCAAT TGCTGATCTC AAGAGAATGAA 13600

Figure 25 (Contd..)

13601 TTCTCGCCTC ACTAATGCCT GAAGAGACCC TCCATCAAGT AATGACACAA CAACCGGGGG ACTCTTCATT
 CCTAGACTGG GCTAGCGACC CTTACTCAGC 13700
 13701 AAATCTTGTA TGTGTCCAGA GCATCACTAG ACTCCTCAAG AACATAACTG CAAGGTTGT CCTGATCCAT
 AGTCCAACC CAATGTTAAA AGGATTATTC 13800
 13801 CATGATGACA GTAAAGAAGA GGACCGGGGA CTGGCGGCAT TCCTCATGGA CAGGCATATT ATAGTACCTA
 GGGCAGCTCA TGAAATCCCTG GATCATAGTG 13900
 13901 TCACAGGGGC AAGAGAGTCT ATTGCAGGCA TGCTGGATAAC CACAAAAGGC TTGATTGAG CCAGCATGAG
 GAAGGGGGGG TAAACCTCTC GAGTGATAAC 14000
 14001 CAGATTGTC C AATTATGACT ATGAACAATT CAGAGCAGGG ATGGTGCTAT TGACAGGAAG AAAGAGAAAT
 GTCCCTCATG ACAAAAGAGTC ATGTTCACTG 14100
 14101 CAGCTGGCGA GAGCTCAAG AAGCCATATG TGGGCGAGGC TAGCTCGAGG ACGGCCTATT TACGGCCTTG
 AGGTCCTCTG TGACTAGAA TCTATGCGAG 14200
 14201 GCCACCTTAT TCGGGCTCAT GAGACATGTG TCATCTGCCA GTGTGGATCA GTCAACTACG GATGGTTTT
 TGTCCCCCTCG GGTTGCCAAC TGGATGATAT 14300
 14301 TGACAAGGAA ACATCATCCT TGAGAGTCCC ATATATTGGT TCTACCACCTG ATGAGAGAAC AGACATGAAG
 CTTGCCTCTG TAAGAGCCCC AAGTCGATCC 14400
 14401 TTGCGATCTG CTGTTAGAAT AGCAACAGTG TACTCATGGG CTTACGGTGA TGATGATAGC TCTTGGAAACG
 AAGCCTGGTT GTGGCTTAGG CAAAGGGCCA 14500
 14501 ATGTGAGCCT GGAGGAGCTA AGGGTGTCA CTCCCACATCTC AACTTCGACT AATTAGCGC ATAGGTTGAG
 GGATCGTAGC ACTCAAGTGA AATACTCAGG 14600
 14601 TACATCCCTT GTCCGAGTGG CGAGGTATAC CACAATCTCC AACGACAATC TCTCATTGT CATATCAGAT
 AAGAAGGTG ATACTAACTT TATATACCAA 14700
 14701 CAAGGAATGC TCCTAGGGT GGGTGTTTA GAAACATTGT TTCGACTCGA GAAAGATACC GGATCATCTA
 ACACGGTATT ACATCTTCAC GTGAAACAG 14800
 14801 ATTGTTGCGT GATCCCAGT ATAGATCATC CCAGGATACC CAGCTCCCGC AAGCTAGAGC TGAGGGCAGA
 GCTATGTACC AACCCATTGA TATATGATAA 14900
 14901 TGCACCTTTA ATTGACAGAG ATGCAACAAAG CCTATACACC CAGAGCCATA GGAGGCACCT TGTGGAATT
 GTTACATGGT CCACACCCCA ACTATATCAC 15000
 15001 'ATTTAGCTA AGTCCACAGC ACTATCTATG ATTGACCTGG TAACAAAATT TGAGAAGGAC CATATGAATG
 AAATTTCAGC TCTCATAGGG GATGACGATA 15100
 15101 TCAATAGTTT CATAACTGAG TTTCTGCTCA TAGAGCCAAG ATTATTCACT ATCTACTTGG GCCAGTGTGC
 GGCCATCAAT TGGCATTG ATGTACATTA 15200
 15201 TCATAGACCA TCAGGAAAT ATCAGATGGG TGAGCTGTT TCATCGTTCC TTTCTAGAAT GAGCAAAGGA
 GTGTTAAGG TGCTTGCTAA TGCTCTAAC 15300
 15301 CACCCAAAGA TCTACAAAGA ATTCTGGCAT TGTGGTATTA TAGAGCCTAT CCATGGCCT TCACTTGATG
 CTCAAACTT GCACACAATC GTGTGCAAA 15400
 15401 TGGTTTACAC ATGCTATATG ACCTACCTCG ACCTGTTGTT GAATGAAGAG TTAGAAGAGT TCACATTCT
 CTTGTGTGAA AGCGACGAGG ATGTTAGTACC 15500
 15501 GGACAGATT CACAACATCC AGGCAAAACA CTTATGTTGTT CTGGCAGATT TGTACTGTCA ACCAGGGACC
 TCCCCACCAA TTGCGAGGTCT AAGACCGGTA 15600
 15601 GAGAAATGTC CAGTTCAAC CGACCATATC AAGGCAGAGG CTATGTTATC TCCAGCAGGA TCTTCGTGGA
 ACATAAATCC AATTATTGTA GACCTATTAC 15700
 15701 CATGCTCTCT GACTTATCTC CGGGCAGGAT CGATCAAACA GATAAGATTG AGAGTTGATC CAGGATTCTAT
 TTTCGACGCC CTCGCTGAGG TAAATGTCAG 15800
 15801 TCAGCCAAG ATCGGCAGCA ACAACATCTC AAATATGAGC ATCAAGGTT TCAGACCCCC ACACGATGAT
 GTTCCAAAT TGCTCAAAGA TATCAACACA 15900
 15901 AGCAAGCACA ATCTTCCAT TTCAAGGGGC AATCTGCCA ATTATGAAAT CCATGTTTC CGCAGAACATCG
 GGGTGAACCT ATCTGCTTGC TACAAAGCTG 16000
 16001 TTGAGATATC AACATTAATT AGGAGATGCC TTGAGCCAGG GGAGGACCGC TTGTTCTTGG GTGAGGGATC
 GGGTTCTATG TTGATCACTT ATAAGGAGAT 16100
 16101 ACTTAAACTA ACAAGTGCT TCTATATAAG TGGGGTTTC GCCAATTCTA GATCTGGTCA AAGGAAATTA
 GCACCCATTC CTCCTCGAAGT TGGGCTTGTG 16200
 16201 GAACACAGAA TGGGAGTAGG TAATATTGTC AAAGTGCTCT TTAACGGGAG GCGCGAAGTC ACGTGGGTAG
 GCAGTGTAGA TTGCTCTAAC TTCACTAGTTA 16300
 16301 GTAATATCCC TACCTCTAGT GTGGGTTTA TCCATTCAAGA TATAGAGACC TTGCTGACA AAGATACTAT
 AGAGAACGTA GAGGAATTGG CAGCCATCTT 16400
 16401 ATCGATGGCT CTGCTCTGG GCAAAATAGG ATCAATACTG GTGATTAAGC TTATGCCTT CAGCGGGGAT
 TTTGTTCAAGG GATTATATAAG TTATGTTAGGG 16500
 16501 TCTCATTATA GAGAAAGTGA CCTGTATAC CCTAGATACA GCAACTTCAT ATCTACTGAA TCTTATTTGG
 TTATGACAGA TCTCAAGGCT AACCGGCTAA 16600
 16601 TGAATCTGA AAAGATTAAG CAGCAGATAA TTGAATCATC TGTGAGGACT TCACCTGGAC TTATAGGTCA
 CATCCTATCC ATTAAGCAAC TAAGCTGCAT 16700
 16701 ACAAGCAATT GTGGGAGACG CAGTTAGTAG AGGTGATATC AACCTACTC TGAAAAAAACT TACACCTATA
 GAGCAAGGTGC TGATCAATTG CGGGTTGGCA 16800
 16801 ATTAACGGAC CTAAGCTGTG CAAAGAATTG ATCCACCATG ATGTTGCCTC AGGGCAAGAT GGATTGCTTA
 ATTCTATATC CATCCTCTAC AGGGAGTTGG 16900
 16901 CAAGATTCAA AGACAACCAA AGAAGTCAAC AAGGGATGTT CCACGCTTAC CCCGTATTGG TAAGTAGCAG
 GCAACGAGAA CTTATATCTA GGATCACCCG 17000

Figure 25(Contd..)

17001 CAAATTTGG GGGCACATTC TTCTTTACTC CGGGAACAGA AAGTTGATAA ATAAGTTTAT CCAGAACATC
 AAGTCCGGCT ATCTGATACT AGACTTACAC 17100
 17101 CAGAATATCT TCGTTAAGAA TCTATCCAAG TCAGAGAAC AGATTATTAT GACGGGGGT TTGAAACGTG
 AGTGGGTTT TAAGTAACA GTCAAGGAGA 17200
 17201 CCAAAGAACG GTATAAGTTA GTCGGATACA GTGCCCTGAT TAAGGACTAA TTGTTGAAC TCCGGAACCC
 TAATCCTGCC CTAGGGTT AGGCATTATT 17300
 17301 TGCAATATAT TAAAGAAAAC TTGAAAATA CGAAGTTCT ATTCCCAGCT TTGTCTGGTg gccggcatgg
 tcccagcctc ctgcgtggcg cggctggc 17400
 17401 aacattccga ggggaccgtc ccctcgtaa tggcgaatgg gacGCGGCCg atccggctgc taacaaagcc
 cggaaaggaaatcgagggtgc tgctgccacc 17500
 17501 gctgagacta aactagata accccctggg gcctctaaac gggctttag gggtttttg ctgaaaggag
 gaactatccggatGCGGC CGCaGGTACC 17600
 17601 CAGCTTTGT TCCCtttagt gagggttaat tTCGAGCTTG GCGTAATCAT GGTCATAGCT GTTTCCTGTG
 TGAAATTGTT ATCCGCTCAC AATTCCACAC 17700
 17701 AACATACGAG CCGGAAGCAT AAAGTGTAAA GCCTGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG
 CGTTGCGCTC ACTGCCCGCT TTTCAGTCGG 17800
 17801 GAAACCTGTC GTGCCAGCTG CATTATGAA TCGGCCAACG CGCGGGGAGA GGCGGTTGC GTATTGGCG
 CTCTCCGCT CTCTCGCTCA CTGAGCTCGT 17900
 17901 GCGCTCGGTG GTTCGGCTGC GGGGAGCGGT ATCAGCTCAC TCAAAGGCCG TAATACGGTT ATCCACAGAA
 TCAGGGGATA ACGCAGGAAA GAACATGTGA 18000
 18001 GCAAAAGGCC AGCAAAAGGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTCCAT AGGCTCCGCC
 CCCCTGACGA GCATCACAAA AATCGACGT 18100
 18101 CAAAGTCAAG GTGCCAACAC CGCAGCAGGAC TATAAAGATA CCAGGCCTT CCCCCGGAA GCTCCCTCGT
 GCGCTCTCTG GTTCGGACCC TGCCGCTTAC 18200
 18201 CGGATACCTG TCCGCCCTTC TCCCTCGGG AAGCGTGGCG CTTTCTCATA GCTCACGCTG TAGGTATCTC
 AGTTCCGGTGT AGGTCCGTTCG CTCCAAGCTG 18300
 18301 GGCTGTGTGC ACGAACCCCC CGTTCAGGCC GACCGCTGGC CCTTATCCGG TAACTATCGT CTTGAGTCCA
 ACCCGGTAAG ACACGACTTA TCGCCACTGG 18400
 18401 CAGCAGCCAC TGGTAACAGG ATTACGAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGT
 GCCTAACTAC GGCTACACTA GAAGGACAGT 18500
 18501 ATTTGGTATC TGCCTCTGC TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTG ATCCGGCAA
 CAAACCCACCG CTGGTAGCGG TGGTTTTTT 18600
 18601 GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTT TCTACGGGT
 CTGACGCTCA GTGGAAACGAA AACTCACGTT 18700
 18701 AAGGGATTTC GGTCATGAGA TTATCAAAAA GGATCTTCAC CTAGATCCTT TAAATTAAA AATGAAGTTT
 TAAATCAATC TAAAGTATAT ATGAGTAAAC 18800
 18801 TTGGTCTGAC AGTTACCAAT GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC
 ATAGTTGCCCT GACTCCCCGT CGTGTAGATA 18900
 18901 ACTACGATCA GGGAGGGCTT ACCATCTGGC CCCAGTGTG CAATGATACC GCGAGACCCA CGCTCACCGG
 CTCCAGATTTC ATCAGCAATA AACCAGCCAG 19000
 19001 CCGGAAGGCC CGAGCGCAGA AGTGGTCTG CAACTTATC CGCCTCCATC CAGTCTATTA ATTGTTGCCG
 GGAAGCTAGA GTAAGTAGTT CGCAGTTAA 19100
 19101 TAGTTGCGC AACGTTGTTG CCATTGCTAC AGGCATCGT GTGTCACGCT CGTCGTTGG TATGGCTTC
 TTCAGCTCCG GTTCCCAACG ATCAAGGCCA 19200
 19201 GTTACATGAT CCCCCATGTT GTGCAAAAAA GCGGTTAGCT CCTTCGGTCC TCCGATCGTT GTCAGAAGTA
 AGTTGGCCGC AGTGTATCA CTCATGGTTA 19300
 19301 TGGCAGCACT GCATAATTCT CTTACTGTCA TGCCATCGT AAGATGCTT TCTGTGACTG GTGAGTACTC
 AACCAAGTCA TTCTGAGAAT AGTGTATGCG 19400
 19401 GCGACCGAGT TGCTCTGCC CGGGCTCAAT ACGGGATAAT ACCGCGCCAC ATAGCAGAAC TTTAAAAGTG
 CTCATCATTG GAAAACGTTC TTCCGGGGCA 19500
 19501 AAACCTCTCAA GGATCTTACG GCTGTTGAGA TCCAGTCGA TGTAACCCAC TCGTGCACCC AACTGATCTT
 CAGCATCTT TACTTCACC AGCGTTCTG 19600
 19601 GGTGAGCAAA AACAGGAAGG CAAAATGCCG CAAAAAAAGGG AATAAGGGCG ACACGGAAAT GTTGAATACT
 CATACTCTTC CTTTTCAAT ATTATTGAAG 19700
 19701 CATTATTCAG GTTATGTC TCATGAGCGG ATACATATT GAATGTATTT AGAAAAATAA ACAAAATAGGG
 GTTCCGCGCA CATTTCCCG AAAAGTGC 19798

Figure 25 (Contd.)

	1	10	1	20	1	30	1	40	1	50	1	60	1
70		80		90		100							
1	ATGaccgtcg	cgccggccgag	cgtccccgcg	gcccgtcccc	tcctcgggaa	gctgccccgg							
	ctgctgtgc	tgggtgtgtt	gtgcctgccc	gccgtgtggG	100								
101	GATCCGTGAC	CCACGAATCC	TATCAGGAGC	TGGTTAAGAA	ACTGGAAGCT	TTAGAGGACG							
	CCGTATTGAC	AGGTTACTCC	CTATCCAGA	AAGAAAAGAT	200								
201	GGTTTTAAC	GAAGAAGAAA	TTACCACAAA	GGGAGCATCC	GCCCAGTCTG	GAGCATCTGC							
	TCAGAGCGGA	GCATCTGCTC	AGAGTGGAGC	AAGCGCCCAA	300								
301	AGTGGAGCGT	CTGCCAGTC	AGGCGCCTCA	GCTCAATCTG	GAACCTCTGG	GCCGAGTGTT							
	CCTAGCGGTA	CTTCTCCAAG	TAGCCGGTCT	AATACACTCC	400								
401	CACGTTCCA	CACCTCCAGT	GGAGCCTCCC	CACCCGCCGA	CGCATCCGAC	TCAGACGCTA							
	AGAGTTATGC	AGACCTGAAG	CACCGCGTGA	GGAACTACCT	500								
501	TTTCACTATC	AAAGAGTGTG	AGTACCCCTGA	ATTGTTCGAT	TTGACCAACC	ATATGCTGAC							
	ACTCTGTGAC	AACATCATG	GTTCAGAAGT	TCTGATAGAT	600								
601	GGGTATGAAG	AAATTAACGA	GCTGCTCTAT	AAACTCAACT	TTTACTTCGA	CCTGCTGCGT							
	GCCAAAGCTGA	ACGATGTCTG	TGCAAACGAT	TAATGCCAGA	700								
701	TCCCCATTCAA	CCTAAAGATA	CGTGCAGACG	AGCTGGATGT	TCTGAAGAAA	CTCGTGTTCG							
	GGTATCGGAA	ACCCCTGGAC	AACATTAAGG	ACAATGTGGG	800								
801	GAAGATGGAG	GATTACATTA	AGAAAAATAA	AACAAACAATC	GCTAACATAA	ATGAGCTTAT							
	CGAGGGGAGC	AAAAAGACCA	TCGACCAGAA	CAAGAATGCC	900								
901	GACAATGAAG	AGGGAAAAAA	GGAACTATAC	CAAGCCCAGT	ATGATTTGAG	CATCTACAAT							
	AACCAACTAG	AGGAAGCTCA	CAACCTCATC	AGCGTACTGG	1000								
1001	AAAAGAGAAT	TGACACCTCG	AAAAAGAATG	AAAACATTAA	GAAACTCCTG	GACAAGATTA							
	ACGAAATTAA	AAACCCaCCT	CCaGCGAATA	GGGAAATAC	1100								
1101	CCCGAATACC	CTGCTGGATA	AGAACAAAAA	GATTGAAGAG	CACGAAGAGA	AAATCAAGGA							
	AATCGCCAAG	ACTATTAAGT	TCAATATAGA	TTCTCTGTT	1200								
1201	ACAGACCCtC	TGGAGCTGGA	ATACTACCTG	CGCGAGAAGA	ATAAGAAGGT	CGACGTGACC							
	CCAAAGAGCC	AAAGACCCAAAC	AAAGTCCGTG	CAGATCCCCA	1300								
1301	AAAGTCCCTA	CCCAAACGGC	ATCGTGTATC	CCCTGCCCT	TACCGACATC	CACAACTCTC							
	TGGCAGCCGA	TAACCGACAAA	AACAGCTATG	GAGACCTGAT	1400								
1401	GAACCCCCAC	ACTAAGGAAA	AGATAAAACGA	GAAGATCATT	ACCGATAATA	AGGAGCGGAA							
	GATTTTATC	AAACAACATCA	AGAAGAAAAAT	CGACCTGGAA	1500								
1501	GAGAAAAATA	TCAATCACAC	CAAAGAGCAA	AACAAAGAAAT	TACTGGAGGA	CTATGAGAAG							
	AGCAAAAGG	ATTATGAGGA	ACTGTTAGAG	AAAGTCTATG	1600								
1601	AAATGAAATT	CAACAACAAT	TTCGATAAGG	ATGTGGTCGA	TAAAATTTTC	AGCGCCCGGT							
	ACACCTACAA	CGTGGAGAAG	CAGCGGTACA	ACAATAAGTT	1700								
1701	CAGCAGCTCC	AAATAACTCGG	TCTACAATGT	GCAGAAGCTG	AAGAAAGCTC	TGAGCTATCT							
	GGAAAGACTAC	TCGCTGAGGA	AAAGGGATTC	TGAGAAGGAT	1800								
1801	TTCAACCACT	ACTACACCC	CAAAACCGGC	CTGGAAGCTG	ACATCAAGAA	ACTCACTGAA							
	GAGATCAAAA	GTTCAGAGAA	TAAGATACTG	GAGAAGAACT	1900								
1901	TCAAGGGACT	AAAGCACTCT	GCAAACGGCT	CCCTGGAAGT	CTCTGACATC	GTGAAACTGC							
	AAGTCCAAAA	GGTGGTGTCTC	ATCAAAAAAA	TCGAGGATCT	2000								
2001	GCGAAAGATC	GAGCTGTTTC	TTAAGAACGC	CCAAGTCAA	GACTCAATCC	ACGTGCCTAA							
	CATTTACAAA	CCGCAGAACAA	AACCAGAAC	ATACTATCTG	2100								
2101	ATCGTGTGA	AGAAGGAGGT	GGATAAGCTG	AAGGAATTCA	TCCCCAAAGT	GAAAGATATG							
	TTAAAGAAAG	AGCAAGCCGT	GCTGAGCAGC	ATAACGCAGC	2200								
2201	CTCTGGTGGC	CGCAAGCGAG	ACAACCGAAG	ATGGCGGGCA	CAGCACCCAC	ACCCCTGTCTC							
	AGTCTGGCGA	AAACAGAGGTG	ACAGAAGAGA	CAGAAGAGAC	2300								
2301	CGAAGAAACA	GTGGGGCACA	CCACTACTGT	GACCACACT	TTGCCCTCTA	CGCAGCCATC							
	TCCCCCAAAA	GAGGTCAAAG	TCGTGGAAAAA	CTCCATTGAA	2400								
2401	CAGAAGTCCA	ACGACAACTC	ACAGGCTCTG	ACGAAGACCG	TCTATCTGAA	GAAACTGGAC							
	GAGTTCTGAA	CCAAAAGCTA	CATCTGCCAT	AAATACATCC	2500								
2501	TCTGTCTAA	CAGCAGCATG	GATCAGAAGC	TGTTGGAGGT	GTACAACCTA	ACGCCCCGAAG							
	AAGAGAACGA	GTAAAATCC	TGTGATCCCT	TAGACCTACT	2600								
2601	GTTTAACATT	CAGAACAAACA	TCCCCGTAT	GTACAGCTTA	TATGATTCCA	TGAATAACGA							
	CCTCCAGCAC	CTGTTCTTCG	AGCTGTACCA	GAAAGAGATG	2700								
2701	ATCTACTATC	TGCATAAGCT	GAAAGAGGAG	AATCACATCA	AAAAGTTGCT	GGAAGAGCAG							
	AAACAGATAA	CTGGGACGTC	CAGCACATCG	TCACCTGGCA	2800								
2801	ACACGACAGT	AAATACCGCC	CAGTCTGCTA	CACACTCCAA	CTCCCAGAAC	CAGCAGAGCA							
	ACGCTTCTAG	CACCAACACC	CAGAATGGGG	TAGCAGTTAG	2900								
2901	TAGCGGCCCT	GCTGTGGTGG	AGGAATCGCA	TGACCCCTC	ACTGTATTAT	CTATTCAA							
	CGACCTAAAA	GGGATTGTGT	CCCTCCTCAA	TTTAGGTAAT	3000								

Figure 26

3001	AAGACCAAGG	TCCCTAACCC	CTTGACTATC	AGCACTACGG	AAATGGAGAA	GTTTTATGAA														
	AACATCTGA	AGAACAAACGA	CACCTATTTT	AACGACGACA	3100															
3101	TAAAGCAGTT	CGTGAAGAGT	AACAGTAAAG	TGATTACCGG	GCTGACAGAA	ACCCAGAAAA														
	ATGCCTTAAA	TGATGAGATC	AAGAAACTGA	AAGACACACT	3200															
3201	CCAGCTCTCC	TTCGATCTGT	ACAACAACTA	CAAACCTAAAG	CTGGACAGAT	TATTCAATAA														
	GAAGAAGGAG	CTTGGGCAAG	ATAAGATGCA	GATTAAGAAG	3300															
3301	CTAACCTTAC	TGAAGGAGCA	GCTCGAGAGC	AAGCTCAACT	CCCTGAATAA	TCCACATAAT														
	GTGCTCCAGA	ACTTTTCCGT	ATTCTTCAAT	AAGAAGAAG	3400															
3401	AAGCAGAGAT	TGCCGAGACG	AAAAATACCC	TCGAAAACAC	TAAGATATTA	CTGAAACACT														
	ATAAAGGCT	GGTGAAGTAT	TACAAACGGAG	AGTCTAGCCC	3500															
3501	ATTGAAGACT	CTTCAGAAG	TGTCAATTCA	AACCGAGGAT	AACTACGCAA	ACCTAGAAAA														
	GTTCAGAGTG	CTGAGCAAAA	TCGACGGCAA	ACTCAATGAT	3600															
3601	AACCTACACC	TCGAAAAAAA	AAAGCTGAGC	TTCCGTCCA	GTGGACTTCA	TCATTTAATT														
	ACCGAATTGA	AAGAAGTTAT	AAAAAACAAA	AACTACACTG	3700															
3701	GGAACAGCCC	ATCTGAAAAT	AATAAAAAGG	TCAACGAGGC	CCTCAAGTCT	TATGAAAATT														
	TCCTTCCAGA	AGCAAAAGTG	ACAACCCTCG	TGACCCCCCCC	3800															
3801	CCAGCCCCAT	GTCACCCCCA	GCCCTCTAAG	CGTGAGAGTG	TCTGGATCAA	GTGGCTCCAC														
	AAAAGAAGAA	ACCCAGATCC	CCACATCAGG	ATCTCTACTG	3900															
3901	ACCGAGTTGC	AGCAGGTCGT	CCAACCTCCAG	AATTATGACG	AGGAAGACGA	CAGCCTCGTG														
	GTTCGGCAA	TCTTCGGCGA	ATCAGAAGAC	AAAGCAGAGT	4000															
4001	ACCTAGACCA	AGTGGTCACC	GGGGAAAGCGA	TTAGTGTAC	TATGGACAAT	ATCCTCAGCG														
	GCTTCGAGAA	CGAGTATGAC	GTGATCTACC	TCAAACCACT	4100															
4101	AGCCGGAGTT	TACAGAAGTC	TCAAAGAACCA	GATCGAAAAG	AACTACCTCA	CCTTTAATCT														
	AAACCTAAAC	GACATCTTGA	ATTCCCGGCT	AAAAAAAGCGG	4200															
4201	AAATACTTCC	TCGACGTACT	GGAGTCGGAT	TTGATGCAGT	TTAAGCACAT	CTCCAGCAAC														
	GAATACATTA	TCGAGGACTC	GTTCAAACCTG	TTAAACTCCG	4300															
4301	AGCAGAAGAA	CACCCCTGCTG	AAAGTCCTACA	AATATATCAA	AGAGTCAGTC	GAGAACGATA														
	TTAAATTCTGC	CCAAGAAGGC	ATAAGCTACT	ACGAAAAGGT	4400															
4401	CCTCGCCAAA	TACAAGGAGC	ATCTGGAGTC	TATCAAAAG	GTCATCAAAG	AAGAGAAAGA														
	GAAATTTCCC	AGTTCTCCCC	CTACAAACGCC	GCCCTCTCCA	4500															
4501	GCCAAGACTG	ATGAACAGAA	AAAAGAGTCT	AAAGTCCTCC	CTTTCCTCAC	TAATATCGAG														
	ACTCTCTACA	ATAAACCTAGT	GAACAAGATT	GACGACTACC	4600															
4601	TGATCAACCT	TAAGCCAAG	ATAAACGACT	GCAATGTCGA	GAAGGATGAG	GCTCATGTTA														
	AGATCACCAA	ACTGTCCGAT	CTGAAAGCCA	TCGACGCCAA	4700															
4701	GATCGACTTA	TTTAAAAAAC	CATACGATT	CGAGGCTATC	AAAAAGCTGA	TCAATGATGA														
	CACCAAGAAA	GATATGCTCG	GCAAGCTGCT	GAGCACGGGT	4800															
4801	CTGGTGCAGA	ACTCCCTAA	CACCATCATA	TCAAAGCTCA	TAGAGGGCAA	GTTCAGAGAC														
	ATGCTGAATA	TTTCACAGCA	TCAGTGCCTC	AAGAAGCAGT	4900															
4901	GCCCCGAAAA	TTCTGGATGC	TTCCGGCACC	TGGATGAGCG	AGAAGAGTGC	AAGTGCCTGC														
	TTAACTATAA	ACAGGAGGGC	GACAAATGTG	TGGAGAACCC	5000															
5001	AAATCCGACG	TGCAACGAGA	ACAACGGTGG	CTGCGATGCC	GACGCGACTT	GTACAGAGGA														
	AGACTCGGGG	AGTTCTCGGA	AAAAAAATCAC	GTGCGACTGC	5100															
5101	ACCAAACCCG	ACAGTTATCC	TCTGTTCCGT	GGGATATTCT	GCTCCTCCAG	CaacgttACT														
	ACTTCCGGCA	CTACCCGTCT	TCTATCTGGT	CACACGTGTT	5200															
5201	TCACGTTGAC	AGGTTTGCTT	GGGACGCTAG	TAACCATGGG	CTTGCTGACT	TAA														
5253																				
	70		80		90		100		10		20		30		40		50		60	

Figure 26 (Contd.)

	10	20	30	40	50	60	70
1	80	90	100				
101	ATGacgctcg	cgcggccgag	cgtcccccg	gcgcgtcccc	tcctcggga	gctgccccgg	
	ctgctgctgc	tggctgttt	gtgcctgccg	gcccgtgtgg	100		
201	101 gatcCGTGAC	CCACGAATCC	TATCAGGAGC	TGGTTAAGAA	ACTGGAAGCT	TTAGAGGACG	
	CCGTATTGAC	AGGTTACTCC	CTATTCAGA	AAGAAAAGAT	200		
301	201 GGTTTAAAC	GAAGAAGAAA	TTACCACAA	GGGAGCATCC	GCCCAGTCTG	GAGCATCTGC	
	TCAGAGCGGA	GCATCTGCTC	ACAGTGGAGC	AAGGCCCAA	300		
401	301 AGTGGACGCT	CTGCCCCAGTC	AGGCGCCCTCA	GCTCAATCTG	GAACCTCTGG	GCCGAGTGGT	
	CCTAGCGGTA	CTTCTCCAAG	TAGCCGGTCT	AATACACTCC	400		
501	401 CACGTTTCAAG	CACCTCCAGT	GGAGCCTCTG	CACCCGCCGA	CGCATCCGAC	TCAGACGCTA	
	AGAGTTATGC	AGACCTGAAG	CACCGCGTGA	GGAACTACCT	500		
601	501 TTTCACTATC	AAAGAGTTGA	AGTACCCCTGA	ATTGTTCGAT	TTGACCAACC	ATATGCTGAC	
	ACTCTGTGAC	AACATACATG	GTTTCAAGTA	TCTGATAGAT	600		
701	601 GGGTATGAAG	AAATTAACGA	GCTGCTCTAT	AAACTCAACT	TTTACTTCGA	CCTGCTGCGT	
	GCCAAGCTGA	ACGATGTCTG	TGCAAACGAT	TACTGCCAGA	700		
801	701 TCCCATTCAA	CCTAAAGATA	CGTGCACG	AGCTGGATGT	TCTGAAGAAA	CTCGTGTTCG	
	GGTATCGGAA	ACCCCTGGAC	AACATTAAGG	ACAATGTGGG	800		
901	801 GAAGATGGAG	GATTACATTA	AGAAAAATAA	AACAACAATC	GCTAACATAA	ATGAGCTTAT	
	CGAGGGGAGC	AAAAAGACCA	TCGACCAGAA	CAAGAATGCC	900		
1001	901 CACAATCAAG	ACGGAAAAAA	CAAACATAC	CAAGCCCAGT	ATGATTTGAG	CATCTACAAT	
	AAGCAACTAG	AGGAAGCTCA	CAACCTCTATC	AGCGTACTGG	1000		
1101	1001 AAAAGAGAAAT	TGACACCTG	AAAAAGAATG	AAAACATTA	GAAACTCCTG	GACAAGATTA	
	ACGAAATTAA	AAACCCAcCt	CCaGCGAATA	GCGGAAATAC	1100		
1201	1101 CCCGAATACC	CTGCTGGATA	AGAACAAAAA	GATTGAAGAG	CACGAAGAGA	AAATCAAGGA	
	AATCGCCAAG	ACTATTAAGT	TCAATATAGA	TTCTCTGTT	1200		
1301	1201 ACAGACCCtC	TGGAGCTGGA	ATACTACCTG	CGCGAGAAGA	ATAAGAAGGT	CGACGTGACC	
	CCAAAGAGCC	AAGACCCAAC	AAAGTCCGTG	CAGATCCCCA	1300		
1401	1301 AAGTCCCTA	CCCAAACGGC	ATCGTGTATC	CCCTGCCTCT	TACCGACATC	CACAACCTCTC	
	TGGCAGCCGA	TAACGACAA	ACAGCTATG	GAGACCTGAT	1400		
1501	1401 GAACCCCCAC	ACTAAGGAAA	AGATAAACGA	GAAGATCATT	ACCGATAATA	AGGAGCGGAA	
	GATTTTATC	AACAAACATCA	AGAAGAAAAAT	CGACCTGGAA	1500		
1601	1501 GAGAAAAATA	TCAATCACAC	CAAAGAGCAA	AAACAGAAAT	TACTGGAGGA	CTATGAGAAG	
	AGCAAAAGG	ATTATGAGGA	ACTGTTAGAG	AAGTTCTATG	1600		
1701	1601 AAATGAAATT	CAACAAACAT	TTCGATAAGG	ATGTGGTCGA	AAAAATTTC	AGCGCCCGGT	
	ACACCTACAA	CGTGGAGAAG	CAGCGGTACA	ACAATAAGTT	1700		
1801	1701 CAGCAGCTCC	AATAACTCGG	TCTACAAATGT	GCAGAAGCTG	AAGAAAGCTC	TGAGCTATCT	
	GGAAGACTAC	TCGCTGAGGA	AAAGGATTT	TGAGAAGGAT	1800		
1901	1801 TTCAACCACT	ACTACACCCt	CAAAACCGGC	CTGGAAGCTG	ACATCAAGAA	ACTCACTGAA	
	GAGATCAAAA	GTTCTGAGAA	TAAGATACTG	GAGAAGAACT	1900		
2001	1901 TCAAGGGACT	AAACGCACTCT	GCAAAACGGCT	CCCTGGAAGT	CTCTGACATC	GTGAAACTGC	
	AAGTCCAAAA	GGTCTGCTC	ATCAAAAAAA	TCGAGGATCT	2000		
2101	2001 GCGAAAGATC	GAGCTGTTTC	TTAAGAACGC	CCAACGTAAA	GACTCAATCC	ACGTGCCTAA	
	CATTTACAAA	CCGAGAACAA	AACCGAACCC	ATACTATCTG	2100		
2201	2101 ATCGTGTGA	AGAAGGAGGT	GGATAAGCTG	AAGGAATTCA	TCCCAAAAGT	GAAAGATATG	
	TTAAAGAAAA	AGCAAGCCGT	GCTGAGCAGC	ATAACCGAGC	2200		
2301	2201 CTCTGGTGGC	CGCAAGCGAG	ACAACCGAG	ATGGCGGGCA	CAGCACCCAC	ACCTGTCTC	
	AGTCTGGCGA	AAACAGAGGTG	ACAGAACAGA	CAGAACAGAC	2300		
2401	2301 CGAAGAAACA	GTGGGGCAC	CCACTACTGT	GACCATCACT	TTGCCCCCTA	CGCAGCCATC	
	TCCCCCAAAA	GAGGTCAAAG	TCTGGAAAAA	CTCCATTGAA	2400		
2501	2401 CAGAAGTCCA	ACGACAAACTC	ACAGGCTCTG	ACGAAGACCG	TCTATCTGAA	GAAACTGGAC	
	GAGTTCTGA	CCAAAAGCTA	CATCTGCCAT	AAATACATCC	2500		
2601	2501 TCGTGTCTAA	CAGCAGCATG	GATCAGAACG	TGTTGGAGGT	GTACAACCTA	ACGCCCGAAG	
	AAGAGAACGA	GTAAAAATCC	TGTGATCCCT	TAGACCTACT	2600		
2701	2601 GTTTAACATT	CAGAACAAACA	TCCCCGCTAT	GTACAGCTTA	TATGATTCCA	TGAATAACGA	
	CCTCCAGCAC	CTGTTCTTCG	AGCTGTACCA	AAAAGAGATG	2700		
2801	2701 ATCTACTATC	TGCTATAAGCT	GAAAGAGGAG	AATCACATCA	AAAAGTTGCT	GGAAGAGCAG	
	AAACAGATAAA	CTGGGACGTC	CAGCACATCG	TCACCTGGCA	2800		
2901	2801 ACACGACAGT	AAATACCGCC	CAGTCTGCTA	CACACTCCAA	CTCCCAGAAC	CAGCAGAGCA	
	ACGCTTCTAG	CACCAACACC	CAGAACGGGG	TAGCAAGTTAG	2900		
3000	2901 TAGCGGGCCT	GCTGTGGTGG	AGGAATCGCA	TGACCCCCCTC	ACTGTATTAT	CTATTCTAAA	
	CGACCTAAAA	GGGATTGTGT	CCCTCCTCAA	TTTAGGTAAT	3000		

Figure 27

3001	AAGACCAAGG	TCCCTAACCC	CTTGACTATC	AGCACTACGG	AAATGGAGAA	GTTTTATGAA		
	AAACATCCTGA	AGAACAACGA	CACCTATTT	AAACGACGACA	3100			
3101	TAAAGCAGT	CGTGAAGAGT	AAACAGTAAAG	TGATTACCGG	GCTGACAGAA	ACCCAGAAAA		
	ATGCTTAAA	TGATGAGATC	AAGAAACTGA	AAGACACACT	3200			
3201	CCAGCTCTCC	TTCGATCTGT	ACAACAAAGTA	CAAACAAAG	CTGGACAGAT	TATTCAATAA		
	GAAGAAGGGAG	CTTGGGCAAG	ATAAGATGCA	GATTAAGAAG	3300			
3301	CTAACCTTAC	TGAAGGGAGCA	GCTCGAGAGC	AAGCTCAACT	CCCTGAATAA	TCCACATAAT		
	GTGCTCCAGA	ACTTTTCCGT	ATTCTTCAT	AAGAAGAAAG	3400			
3401	AAGCAGAGAT	TGCCGAGACG	GAAAATACCC	TCGAAAACAC	TAAGATATTA	CTGAAACACT		
	ATAAAGGGCT	GGTGAAGTAT	TACAACCGGAG	AGTCTAGCCC	3500			
3501	ATTGAAGACT	CTTCAGAAG	TGTCAATTC	AACCGAGGAT	AACTACGCAA	ACCTAGAAAA		
	GTTCAGAGTG	CTGAGCAAAA	TCGACGGCA	ACTCAATGAT	3600			
3601	AAACCTACACC	TCGAAAAAAA	AAAGCTGAGC	TTCCGTCCA	GTGGACTTCA	TCATTTAATT		
	ACCGAATTGA	AAAAGAGTTAT	CAAAAACAAA	AACTACACTG	3700			
3701	GGAACAGCCC	ATCTGAAAAT	AATAAAAAGG	TCAACGAGGC	CCTCAAGTCT	TATGAAAATT		
	TCCTTCCAGA	AGCAAAAGTG	ACAACCGTCG	TGACCCCCCCC	3800			
3801	CCAGCCCCAT	GTCACCCCCA	GCCCTCTAAG	CGTAGAGTGT	TCTGGATCAA	GTGGCTCCAC		
	AAAAGAAGAA	ACCCAGATCC	CCACATCAGG	ATCTCTACTG	3900			
3901	ACCGAGTTCG	AGCAGGTCGT	CCAACCTCAG	AATTATGACG	AGGAAGACGA	CAGCCTCGTG		
	GTTCAGGCAAA	TCTTCGGCGA	ATCAGAAGAC	AAACGACGAGT	4000			
4001	ACCTAGACCA	AGTGGTCACC	GGGGAAAGCGA	TTAGTGTAC	TATGGACAAT	ATCCTCAGCG		
	GCTTCGAGAA	CGAGTATGAC	GTGATCTACC	TCAAACACT	4100			
4101	ACCCGGAGTT	TACAGAAGTC	TCAAGAACCA	GATCGAAAAG	AAACATCTCA	CCTTTAATCT		
	AAACCTAAAC	GACATCTTGA	ATTCCCGGCT	GAAAAAGCGG	4200			
4201	AAATACTTCC	TCGACGTACT	GGAGTCGGAT	TTGATGCGAT	TTAAGCACAT	CTCCAGCAAC		
	GAATACATTA	TCGAGGACTC	GTTCAAACGT	TTAAACTCCG	4300			
4301	AGCAGAAGAA	CACCCCTGCTG	AAAGTCCTACA	AATATATCAA	AGAGTCAGTC	GAGAACGATA		
	TTAAATTCGC	CCAAGAAGGC	ATAAGCTACT	ACGAAAAGGT	4400			
4401	CCTCGCCAAA	TACAAGGACG	ATCTGGAGTC	TATCAAAAG	GTCATCAAAG	AAGAGAAAGA		
	GAAATTCTCC	AGTTCTCCCC	CTACAACGCC	GCCCTCTCCA	4500			
4501	GCCAAGACTG	ATGAACAGAA	AAAAGACTCT	AAAGTCCTCC	CTTTCCTCAC	TAATATCGAG		
	ACTCTCTACA	ATAACCTAGT	GAACAAAGATT	GACGACTACC	4600			
4601	TGATCAACCT	TAAAGCCAAG	ATAAACGACT	GCAATGTCGA	GAAGGATGAG	GCTCATGTTA		
	AGATCACCAA	ACTGTCCGAT	CTGAAAGCCA	TCGACGACAA	4700			
4701	GATCGACTTA	TTTAAAAAAC	CATACGATTT	CGAGGCTATC	AAAAAGCTGA	TCAATGATGA		
	CACCAAGAAA	GATATGCTCG	GCAAGCTGCT	GAGCACGGGT	4800			
4801	CTGGTGAGA	ACTTCCTAA	CACCATCAT	TCAAAGCTCA	TAGAGGGCAA	GTTCCAAGAC		
	ATGCTGAATA	TTTCACAGCA	TCAGTGCCTC	AAGAAGCAGT	4900			
4901	CCCCCGAAAA	TTCGGATGC	TTCCGGCACC	TGGATGAGCG	AGAAGAGTGC	AAGTGCCTGC		
	TTAACTATAA	ACAGGAGGGC	GACAAATGTG	TGGAGAACCC	5000			
5001	AAATCCGACG	TGCAACGAGA	ACAACGGTGG	CTGCGATGCC	GACCGCAGTT	GTACAGAGGA		
	AGACTCGGGG	AGTTCTCGGA	AAAAAATCAC	GTGCGAGTGC	5100			
5101	ACCAAACCCG	ACAGTTATCC	TCTGTTCGAT	GGGATATTCT	GCTCCTCCAG	CAACGTTAG		
5160								
		10	20	30	40	50	60	
70		80	90	100				

Figure 27 (Contd.)

	10		20		30		40		50		60		70
	80		90		100								
1	ATGaccgtcg	cggggccgag	cgtcccccg	gcgcgtcccc	tcctcggga		gctgccccgg						
	ctgctgtgc	tgggtgtgtt	gtgcctgccc	gcgcgtgtggG	100								
101	GATCCGTGAC	CCACGAATCC	TATCAGGAGC	TGGTTAACGAA	ACTGGAAGCT		TTAGAGGACG						
	CCGTATTGAC	AGTTTACTCC	CTATTCCAGA	AAGAAAAGAT	200								
201	GGTTTAAAC	GAAGAAGAAA	TTACCACAAA	GGGAGCATCC	GCCCCAGTCTG		GAGCATCTGC						
	TCAGAGCGGA	GCATCTGCTC	AGAGTGGAGC	AAGGCCCAA	300								
301	AGTGGAGCGT	CTGCCCAGTC	AGGCAGCTCA	GCTCAATCTG	GAACCTCTGG		GCCGAGTGGT						
	CCTAGCGGTA	CTTCTCCAAG	TAGCCGGTCT	AATACACTCC	400								
401	CACGTTCAA	CACCTCCAGT	GGAGCCTCCC	CACCCGCCGA	CGCATCCGAC		TCAGACGCTA						
	AGAGTTATGC	AGACCTGAAG	CACCGCGTGA	GGAACTACCT	500								
501	TTTCACTATC	AAAGAGTTGA	AGTACCCCTGA	ATTGTTCGAT	TTGACCAACC		ATATGCTGAC						
	ACTCTGTGAC	AACATACATG	GTTCAGTGA	TCTGATAAGAT	600								
601	GGGTATGAAG	AAATTAACGA	GCTGCTCTAT	AAACTCAACT	TTTACTTCGA		CCTGCTGCGT						
	GCCAAGCTGA	ACGATGTCTG	TGCAAACGAT	TACTGCCAGA	700								
701	TCCCATTCAA	CCTAAAGATA	CGTGCAGACG	AGCTGGATGT	TCTGAAGAAA		CTCGTTCG						
	GGTATCGGAA	ACCCCTGGAC	ACATTAAGG	ACAATGTGGG	800								
801	GAAGATGGAG	GATTACATTA	AGAAAAATAA	AAACAAATC	GCTAACATAA		ATGAGCTTAT						
	CGAGGGAGC	AAAAGACCA	TCGACCGAGA	CAAGAATGCC	900								
901	GACAATGAAG	AGGGAAAAAA	GAAACTATAC	CAAGCCCAGT	ATGATTTGAG		CATCTACAAT						
	AAGCAACTAG	AGGAAGCTCA	CAACCTCATC	AGCGTACTGG	1000								
1001	AAAAGAGAAT	TGACACCCCTG	AAAAAGAATG	AAAACATTAA	GAAACTCCTG		GACAAGATTA						
	ACGAAATTAA	AAACCCcAcCt	CCaGCGAATA	GCGGAAATAC	1100								
1101	CCCGAATACC	CTGCTGGATA	AGAACAAAAA	GATTGAAGAG	CACGAAGAGA		AAATCAAGGA						
	AATCGCCAAG	ACTATTAAGT	TCATATAGA	TTCTCTGTTC	1200								
1201	ACAGACCCtC	TGGAGCTGGA	ATACTACCTG	CGCGAGAAGA	ATAAGAAGGT		CGACGTGACC						
	CCAAAGAGCC	AAAGACCCAAAC	AAAGTCCGTG	CAGATCCCCA	1300								
1301	AAAGTCCCTA	CCCAAACGGC	ATCGTGTATC	CCCTGCCCT	TACCGACATC		CACAACTCTC						
	TGGCAGCCGA	TAACGACAAA	AAACAGCTATG	GAGACCTGAT	1400								
1401	GAACCCCCAC	ACTAAGGAAA	AGATAAAACGA	GAAGATCATT	ACCGATAATA		AGGAGCGGAA						
	GATTTTTATC	AACACATCA	AGAAGAAAAT	CGACCTGGAA	1500								
1501	GAGAAAAATA	TCAATCACAC	CAAAGAGCAA	AAACAGAAAAT	TACTGGAGGA		CTATGAGAAG						
	AGCAAAAGG	ATTATGAGGA	ACTGTTAGAG	AAGTTCTATG	1600								
1601	AAATGAAATT	CAACAACAA	TTCGATAAGG	ATGTGGTCGA	TAAAATTTTC		AGCGCCCGT						
	ACACCTACAA	CGTGGAGAAG	CAGCGGTACA	ACAATAAGTT	1700								
1701	CAGCAGCTCC	AAATAACTCGG	TCTACAAATGT	GCAGAAGCTG	AAGAAAGCTC		TGAGCTATCT						
	GGAAGACTAC	TCGCTGAGGA	AAAGGATTTC	TGAGAAGGGAT	1800								
1801	TTCAACCACT	ACTACACCCCT	CAAAACGGC	CTGGAAGCTG	ACATCAAGAA		ACTCACTGAA						
	GAGATCAAAA	GTTCTGAGAA	TAAGATACTG	GAGAAGAACT	1900								
1901	TCAAGGACT	AACGCACCT	GCAAACGCT	CCCTGGAACT	2000		CTCTGACATC						
	AACTCCAAA	GGTGTGCTC	ATCAAAAAAA	TCGAGGATCT			GTGAAACTC						
2001	GCGAAAGATC	GAGCTGTTTC	TTAAGAACGC	CCAAGTGAAA									
	CATTTACAAA	CCGAGAAACA	AACCAGAAC	ATACTATCTG	2100								
2101	ATCGTGTGAA	AGAAGGGATG	GGATAAGCTG	AAGGAATTCA	TCCCAAAAGT		GAAAGATATG						
	TTAAAGAAAG	ACCAAGCCGT	GCTGAGGAGC	ATAACCCAGC	2200								
2201	CTCTGGTGGC	CGCAAGCGAG	ACAACCGAAG	ATGGCGGGCA	CAGCACCCAC		ACCCGTCTC						
	AGTCTGGCGA	AACAGAGGTG	ACAGAAGAGA	CAGAAGAGAC	2300								
2301	CGAAGAAACA	GTGGGGCACA	CCACTACTGT	GACCATCACT	TTGCCCCCTA		CGCAGCCATC						
	TCCCCAAAAA	GAGGTCAAAG	TCGTGGAAAA	CTCCATTGAA	2400								
2401	CAGAAGTCCA	ACGACAAACTC	ACAGGCTCTG	ACGAAGACCG	TCTATCTGAA		GAAACTGGAC						
	GAGTTCTGTA	CCAAAAGCTA	CATCTGCCAT	AAATACATCC	2500								
2501	TCGTGTCTAA	CAGCAGCATG	GATCAGAACGC	TGTTGGAGGT	GTACAAACCTA		ACGCCGAAG						
	AAGAGAACGA	GTTAAAATCC	TGTGATCCCT	TAGACCTACT	2600								
2601	GTTTAACATT	CAGAACAAACA	TCCCCGCTAT	GTACAGCTTA	TATGATTCCA		TGAATAACGA						
	CCTCCAGCAC	CTGTTCTTCG	AGCTGTACCA	GAAAGAGATG	2700								
2701	ATCTACTATC	TGCTATAAGCT	GAAAGAGGAG	AATCACATCA	AAAAGTTGCT		GGAAGAGCAG						
	AAACAGATAA	CTGGGACGTC	CAGCACATCG	TCACCTGGCA	2800								
2801	ACACGACAGT	AAATACCGGCC	CAGTCTGCTA	CACACTCAA	CTCCCAGAAC		CAGCAGAGCA						
	ACGCTCTAG	CACCAACACC	CAGAATGGGG	TAGCAGTTAG	2900								
2901	TAGCGGCCCT	GCTGTGGTGG	AGGAATCGCA	TGACCCCTC	ACTGTATTAT		CTATTTCAA						
	CGACCTAAAA	GGGATTGTGT	CCCTCCTCAA	TTTAGGTAAT	3000								

Figure 28

3001	AAGACCAAGG	TCCCTAACCC	CTTGACTATC	AGCACTACGG	AAATGGAGAA	GTTTTATGAA											
	AACATCCTGA	AGAACAAACGA	CACCTATTT	AAACGACGACA	3100												
3101	TAAAGCAGTT	CGTGAAGAGT	AACAGTAAAG	TGATTACCGG	GCTGACAGAA	ACCCAGAAAA											
	ATGCTTTAAA	TGATGAGATC	AAGAAACTGA	AAGACACACT	3200												
3201	CCAGCTCTCC	TTCGATCTGT	ACAACAAGTA	CAAACCTAAAG	CTGGACAGAT	TATTCAATAA											
	GAAGAAGGAG	CTTGGGCAAG	ATAAGATGCA	GATTAAGAAG	3300												
3301	CTAACCTTAC	TGAAGGAGCA	GCTCGAGAGC	AAGCTCAACT	CCCTGAATAA	TCCACATAAT											
	GTGCTCCAGA	ACTTTCCGT	ATTCTTCAT	AAGAAGAAG	3400												
3401	AAGCAGAGAT	TGCCGAGACG	GAAAATACCC	TCGAAAACAC	TAAGATATTA	CTGAAACACT											
	ATAAAGGGCT	GGTGAAGTAT	TACAACGGAG	AGTCTAGCCC	3500												
3501	ATTGAAGACT	CTTTCAGAAG	TGTCAATTCA	AACCGAGGAT	AACTACGCAA	ACCTAGAAAA											
	GTTCAGAGTG	CTGAGCAAAA	TCGACGGCAA	ACTCAATGAT	3600												
3601	AACCTACACC	TCGGAAAAAA	AAAGCTGAGC	TTCCTGTCCA	GTGGACTTCA	TCATTTAATT											
	ACCGAATTGA	AAGAAGTTAT	CAAAAACAAA	AACTACACTG	3700												
3701	GGAACAGCCC	ATCTGAAAAT	AATAAAAAGG	TCAACGAGGC	CCTCAAGTCT	TATGAAAATT											
	TCCTTCCAGA	AGCAAAAGTG	ACAACCGTCG	TGACCCCCC	3800												
3801	CCAGCCCGAT	GTCACCCCCA	GCCCTCTAAG	CGTGAGAGTG	TCTGGATCAA	GTGGCTCCAC											
	AAAAGAAGAA	ACCCAGATCC	CCACATCAGG	ATCTCTACTG	3900												
3901	ACCGAGTGC	AGCAGGTCGT	CCAACCTCAG	AATTATGACG	AGGAAGACGA	CAGCCTCGTG											
	GTTTTGCCAA	TCTTCGGCGA	ATCAGAAAGAC	AACGACGGAGT	4000												
4001	ACCTAGACCA	AGTGgtcacC	aacgttACTA	CTTCCGGCAC	TACCCGTCTT	CTATCTGGTC											
	ACACGTGTTT	CACGTTGACA	GGTTTGCTTG	GGACGCTAGT	4100												
4101		AACCATGGGC		TTGCTGACTT		AA											
4122																	
	70		80	10		90	20		100	30		40		50		60	

Figure 28(Contd.)

	10	20	30	40	50	60	70	
			80	90	100			
1	ATGaccgtcg	cgccggccgag	cgtgcccgcg	gcccgtcccc	tcctcgggga	gctgccccgg		
	ctgctgctgc	tggctgtt	gtccctgcg	gcccgttgG	100			
101	GATCCGTGAC	CCACGAATCC	TATCAGGAGC	TGGTTAAGAA	ACTGGAAGCT	TTAGAGGACG		
	CCGTATTGAC	AGGTTACTCC	CTATTCCAGA	AAGAAAAGAT	200			
201	GGTTTAAAC	GAAGAAGAAA	TTACCACAAA	GGGAGCATCC	GCCCAGTCG	GAGCATCTGC		
	TCAGAGCGGA	GCATCTGCTC	AGAGTGGAGC	AAGGCCCAA	300			
301	AGTGGAGCGT	CTGCCAGTC	AGGGCCCTCA	GCTCAATCTG	GAACCTCTGG	GCCGAGTGGT		
	CCTAGCGGTA	CTTCTCCAAG	TAGCCGGTCT	AATAACTCTC	400			
401	CACGTTCCAA	CACCTCAGT	GGAGCCTCC	CACCCGCCGA	CGCATCCGAC	TCAGACGCTA		
	AGAGTTATGC	AGACCTGAAG	CACCGCGTGA	GGAACTACCT	500			
501	TTTCACTATC	AAAGAGTTGA	AGTACCCCTGA	ATTGTTCGAT	TTGACCAACC	ATATGCTGAC		
	ACTCTGTGAC	AACATACATG	GTTCAGTA	TCTGATAGAT	600			
601	GGGTATGAAG	AAATAAACGA	GCTGCTCTAT	AAACTCAACT	TTTACTTCGA	CCTGCTGCGT		
	GCCAAGCTGA	ACGATGCTTG	TCGAAACGAT	TACTGCCAGA	700			
701	TCCCATTCAA	CCTAAAGATA	CGTGCAGACG	AGCTGGATGT	TCTGAAGAAA	CTCGTGTTCG		
	GGTATCGGAA	ACCCCTGGAC	AACATTAAGG	ACAATGTGGG	800			
801	GAAGATGGAG	GATTACATTA	AGAAAAATAA	AAACACAATC	GCTAACATAA	ATGAGCTTAT		
	CGAGGGGAGC	AAAAAGACCA	TCGACCAGAA	CAAGAATGCC	900			
901	GACAATGAAG	AGGGAAAAAA	GAAACTATAC	CAAGCCAGT	ATGATTGAG	CATCTACAAT		
	AAGCAACTAG	AGGAAGCTCA	CAACCTCATC	AGCGTACTGG	1000			
1001	AAAAGAGAAT	TGACACCCCTG	AAAAAGAAAT	AAAACATTAA	GAAACTCCTG	GACAAGATTA		
	ACGAAATTAA	AAACCCCaCCT	CcaGCGATA	GCGGAAATAC	1100			
1101	CCCGAATACC	CTGCTGGATA	AGAACAAAAA	GATTGAAGAG	CACGAAGAGA	AAATCAAGGA		
	AATCGCCAAG	ACTATTAAGT	TCAATAAAGA	TTCTCTGTT	1200			
1201	ACAGACCCCTC	TGGAGCTGGA	ATACTACCTG	CGCGAGAAGA	ATAAGAAGGT	CGACGTGACC		
	CCAAAGAGCC	AAGACCAAC	AAAGTCCGTG	CAGATCCCCA	1300			
1301	AAGTGCCTA	CCCAAACGGC	ATCGTGTATC	CCCTGCCTCT	TACCGACATC	CACAACCTCTC		
	TGGCAGCCGA	TAACGACAAA	AAACAGCTATG	GAGACCTGTAT	1400			
1401	GAACCCCCAC	ACTAAGGAAA	AGATAAACGA	GAAGATCATT	ACCGATAATA	AGGAGCGGAA		
	GATTTTATC	AACAAACATCA	AGAAGAAAAAT	CGACCTGGAA	1500			
1501	GAGAAAAATA	TCAATCACAC	CAAAGAGCAA	AAACAGAAAT	TACTGGAGGA	CTATGAGAAG		
	AGCAAAAGG	ATTATGAGGA	ACTGTTAGAG	AAAGTTCTATG	1600			
1601	AAATGAAATT	CAACAAACAT	TTCGATAAGG	ATGTTGTCGA	TAAAATTTTC	AGCGCCCGGT		
	ACACCTACAA	CGTGGAGAAG	CAGCGGTACA	ACAATAAGTT	1700			
1701	CAGCAGCTCC	AATAACTCGG	TCTACAATGT	GCAGAAGCTG	AAGAAAGCTC	TGAGCTATCT		
	GGAAGACTAC	TCGCTGAGGA	AAGGGATTTC	TGAGAAGGAT	1800			
1801	TTCAACCACT	ACTACACCCCT	CAAAACCGGC	CTGGAAGCTG	ACATCAAGAA	ACTCACTGAA		
	GAGATCAAAA	GTTCTGAGAA	TAAGATACTG	GAGAAGAACT	1900			
1901	TCAAGGGACT	AACGCACTCT	GCAAACGGCT	CCCTGGAAGT	CTCTGACATC	GTGAAACTG		
	AAGTCCAAA	GGTGCCTGCTC	ATCAAAAAAA	TCGAGGATCT	2000			
2001	GCGAAAGATC	GAGCTGTTTC	TTAAGAACGC	CCAACGTAAA	GACTCAATCC	ACGTGCCTAA		
	CATTTACAAA	CCGCGAGAAC	AACCAGAAC	ATACTATCTG	2100			
2101	ATCGTGTGA	AGAAGGAGGT	GGATAAGCTG	AAGGAATTCA	TCCCAAAAGT	GAAAGATATG		
	TTAAAGAAAG	AGCAAGCCGT	GCTGAGCAGC	ATAACCGAGC	2200			
2201	CTCTGGTGGC	CGCAAGCGAG	ACAACCGAAG	ATGGCGGGCA	CAGCACCCAC	ACCCCTGTCTC		
	AGTCTGGCGA	AACAGAGGTG	ACAGAAGAGA	CAGAAGAGAC	2300			
2301	CGAAGAAACAA	GTGGGGCACA	CCACTACTGT	GACCACACT	TTGCCCCCTA	CGCAGCCATC		
	TCCCCAAA	GAGGTCAAAG	TCGTGGAAAA	CTCCATTGAA	2400			
2401	CAGAAGTCCA	ACGACAACTC	ACAGGCTCTG	ACGAAGACCG	TCTATCTGAA	GAAACTGGAC		
	GAGTTCTGA	CCAAAAGCTA	CATCTGCCAT	AAATACATCC	2500			
2501	TCGTGTCTAA	CAGCAGCATG	GATCAGAACG	TGTTGGAGGT	GTACAACCTA	ACGCCCGAAG		
	AAGAGAACGA	GTAAAATCC	TGTGATCCCT	TAGACCTACT	2600			
2601	GTTTAACATT	CAGAACACA	TCCCCGCTAT	GTACAGCTTA	TATGATTCCA	TGAATAACGA		
	CCTCCAGCAC	CTGTTCTTCG	AGCTGTACCA	GAAAGAGATG	2700			
2701	ATCTACTATC	TGCATAAGT	GAAAGAGGAG	AATCACATCA	AAAAGTTGCT	GGAAGAGCAG		
	AAACAGATAA	CTGGGACGTC	CAGCACATCG	TCACCTGGCA	2800			
2801	ACACGACAGT	AAATACCGCC	CAGTCTGCTA	CACACTCCAA	CTCCCAGAAC	CAGCAGAGCA		
	ACGCTTCTAG	CACCAACACC	CAGAATGGGG	TAGCAGTTAG	2900			
2901	TAGCGGGCCCT	GCTGTTGGTGG	AGGAATCGCA	TGACCCCTC	ACTGTATTAT	CTATTCAAA		
	CGACCTAAA	GGGATTGTGT	CCCTCCTCAA	TTTAGGTAAT	3000			

Figure 29

3001	AAGACCAAGG	TCCCTAACCC	CTTGACTATC	AGCACTACGG	AAATGGAGAA	GTTTTATGAA											
	AACATCCTGA	AGAACAAACGA	CACCTATTT	AAACGACGACA	3100												
3101	TAAAGCAGTT	CGTGAAAGAGT	AACAGTAAAG	TGATTACCGG	GCTGACAGAA	ACCCAGAAAA											
	ATGCTTTAAA	TGATGAGATC	AAGAAACTGA	AAGACACACT	3200												
3201	CCAGCTCTCC	TTCGATCTGT	ACAACAAAGTA	CAAACATAAG	CTGGACAGAT	TATTCAATAA											
	GAAGAAGGAG	CTTGGGCAAG	ATAAGATGCA	GATTAAGAAG	3300												
3301	CTAACCTTAC	TGAAGGAGCA	GCTCGAGAGC	AAGCTCAACT	CCCTGAATAA	TCCACATAAT											
	GTGCTCCAGA	ACTTTCCGT	ATTCTTCAT	AAGAAGAAAG	3400												
3401	AAGCAGAGAT	TGCCGAGACG	AAAAATACCC	TCGAAAACAC	TAAGATATTA	CTGAAACACT											
	ATAAAGGCT	GGTGAAGTAT	TACAACGGAG	AGTCTAGCCC	3500												
3501	ATTGAAGACT	CTTCAGAAG	TGTCAATTCA	AACCGAGGAT	AACTACGCAA	ACCTAGAAAA											
	GTTCAGAGTG	CTGAGCAAAA	TCGACGGCAA	ACTCAATGAT	3600												
3601	AACCTACACC	TCGGAAAAAA	AAAGCTGAGC	TTCCTGTCCA	GTGGACTTCA	TCATTTAATT											
	ACCGAATTGA	AAGAAGTTAT	AAAAAACAAA	AACTACACTG	3700												
3701	GGAACAGCCC	ATCTGAAAAT	AATAAAAAGG	TCAACGAGGC	CCTCAAGTCT	TATGAAAATT											
	TCCTTCCAGA	AGCAAAAGTG	ACAACCGTCG	TGACCCCCCC	3800												
3801	CCAGCCCCAT	GTCACCCCCC	GCCCTCTAAG	CGTGAGAGTG	TCTGGATCAA	GTGGCTCCAC											
	AAAAGAAGAA	ACCCAGATCC	CCACATCAGG	ATCTCTACTG	3900												
3901	ACCGAGTTGC	AGCAGGTCGT	CCAACTCCAG	AATTATGACG	AGGAAGACGA	CAGCCTCGTG											
	GTTTTGCCAA	TCTTCGGCGA	ATCAGAAGAC	AAACGACGAGT	4000												
4001		ACCTAGACCA		AGTGGTCACC		GGGGAATAA											
4029																	
	70		80		10		20		30		40		50		60		

Figure 29 (Contd.)

		10		20		30		40		50		60		70	
		80		90		100									

```

1 ATGaccgtcg cgcggccgag cgtccccg  ggcgtcccc  tcctcgggg  gctccccgg
ctgctgtgc tggtgtgtt gtgcctgccg gccgtgtgg  100
101 GATCCGTGGT CACCGGGGAA GCGATTAGTG TCACTATGGA CAATATCCTC AGCGGCTTCG
AGAACGAGTA TGACGTGATC TACCTCAAC CACTAGCCGG 200
201 AGTTTACAGA AGTCTCAAGA ACCAGATCGA AAAGAACATC TTCACCTTA ATCTAACCT
AAACGACATC TTGAATTCCC GGCTGAAAAA GCGGAATAC 300
301 TTCCCTCGACG TACTGGAGTC GGATTTGATG CAGTTAACG ACATCTCCAG CAACGAATAC
ATTATCGAGG ACTCGTTCAA ACTGTTAACC TCCGAGCAGA 400
401 AGAACACCCCT GCTGAAGTCC TACAAATATA TCAAAGAGTC AGTCGAGAAC GATATTAAAT
TCGCCAAGA AGGCATAAGC TACTACGAA AGGTCCCTCGC 500
501 CAAATACAAG GACGATCTGG AGTCTATCAA AAAGGTCATC AAAGAAGAGA AAGAGAAATT
TCCCCAGTTCT CCCCTACAAA CGCCGCCCTC TCCAGCCAAG 600
601 ACTGATGAAC AGAAAAAAGA GTCTAAGTC CTCCCTTTCC TCACTAATAT CGAGACTCTC
TACAATAACC TAGTGAACAA GATTGACGAC TACCTGATCA 700
701 ACCTTAAAGC CAAGATAAAC GACTGCAATG TCGAGAAGGA TGAGGCTCAT GTTAAGATCA
CCAAACTGTC CGATCTGAAA GCCATCGACG ACAAGATCGA 800
801 CTTATTTAAA AACCCATACG ATTCGAGGC TATCAAAAG CTGATCAATG ATGACACCAA
GAAAGATATG CTCGGCAAGC TGCTGAGCAC GGGTCTGGTG 900
901 CAGAACTTCC CTAACACCAT CATATCAAAG CTCATAGAGG GCAAGTTCCA AGACATGCTG
AATATTCAC AGCATCAGTG CGTCAAGAAG CAGTGCCCG 1000
1001 AAAATTCTGG ATGCTTCCGG CACCTGGATG AGCGAGAAGA GTGCAAGTGC CTGCTTAACT
ATAAACAGGA GGGCGACAAA TGTGTGGAGA ACCCAATCC 1100
1101 GACGTGCAAC GAGAACAAACG GTGGCTGCCA TGCCGACGCG ACTTGTACAG AGGAAGACTC
GGGGAGTTCT CGGAAAAAAA TCACGTGCGA GTGCACCAAA 1200
1201 CCCGACAGTT ATCCTCTGTT CGATGGGATA TTCTGCTCCT CCAGCaacgt tACTACTTCC
GGCACTACCC GTCTTCTATC TGGTCACACG TGTTCACGT 1300
1301 TGACAGGTTT GCTTGGGACG CTAGTAACCA TGGGCTTGCT GACTTAA
1347

```

		10		20		30		40		50		60		
		70		80		90		100						

Figure 30

	10	20	30	40	50	60	70
1	80	90	100				
1	ATGaccgtcg	cgcggccgag	cgtcccccg	gcgcgtcccc	tcctcgggga	gctgccccgg	
	ctgctgtgc	tggtgtgtt	gtgcctgccc	gccgtgtggg	100		
101	gatcCGTGGT	CACCGGGGAA	GCGATTAGTG	TCACTATGGA	CAATATCCTC	AGCGGCTTCG	
	AGAACGAGTA	TGACGTGATC	TACCTCAAAC	CACTAGCCGG	200		
201	AGTTTACAGA	AGTCTCAAGA	AGCAGATCGA	AAAGAACATC	TTCACCTTTA	ATCTAAACCT	
	AAACGACATC	TTGAATTCCC	GGCTGAAAAA	GCGGAATAAC	300		
301	TTCCCTCGACG	TACTGGAGTC	GGATTTGATG	CAGTTTAAGC	ACATCTCCAG	CAACGAATAAC	
	ATTATCGAGG	ACTCGTTCAA	ACTGTTAAC	TCCGAGCAGA	400		
401	AGAACACCCCT	GCTGAAGTCC	TACAAATATA	TCAAAGAGTC	AGTCGAGAAC	GATATTAAAT	
	TCGCCCAAGA	AGGCATAAGC	TACTACGAAA	AGGTCCCTCGC	500		
501	CAAATACAAG	GACGATCTGG	AGTCTATCAA	AAAGGTCATC	AAAGAAGAGA	AAGAGAAATT	
	TCCCAGTTCT	CCCCCTACAA	CGCCGCCCTC	TCCAGCCAAG	600		
601	ACTGATGAAC	AGAAAAAAGA	GTCTAAGTC	CTCCCTTTCC	TCACTAATAT	CGAGACTCTC	
	TACAATAACC	TAGTGAACAA	GATTGACGAC	TACCTGATCA	700		
701	ACCTTAAAGC	CAAGATAAAC	GACTGCAATG	TCGAGAAGGA	TGAGGCTCAT	GTAAAGATCA	
	CCAAACTGTC	CGATCTGAAA	GCCATCGACG	ACAAGATCGA	800		
801	CTTATTTAAA	AACCCATACG	ATTCGAGGC	TATCAAAAG	CTGATCAATG	ATGACACCAA	
	GAAAGATATG	CTCGGCAAGC	TGCTGAGCAC	GGGTCTGGTG	900		
901	CAGAACTTCC	CTAACACCAT	CATATCAAAG	CTCATAGAGG	GCAAGTTCCA	AGACATGCTG	
	AATATTTAC	AGCATCAGTG	CGTCAAGAAG	CAGTGCCCCG	1000		
1001	AAAATTCTGG	ATGCTTCCGG	CACCTGGATG	AGCGAGAAGA	GTGCAAGTGC	CTGCTTAACT	
	ATAAACAGGA	GGGCGACAAA	TGTGTGGAGA	ACCCAAATCC	1100		
1101	GACGTGCAAC	GAGAACAAACG	GTGGCTGCGA	TGCCGACGCG	ACTTGTACAG	AGGAAGACTC	
	GGGGAGTTCT	CGGAAAAAAA	TCACGTGCGA	GTGCACCAAA	1200		
1201	CCCGACAGTT	ATCCTCTGTT	CGATGGGATA	TTCTGCTCCT	CCAGCAACGT	TTAG	
1254							
	70	80	90	100	10	20	30

Figure 31

Figure 32

3001	TCTAAaCGtA	AGAAgCTGGA	aGAGGACATC	AATAAgctgA	AgAAGACaCT	gCAaCTGagc																																																																																																																																								
	TTCGACcTGT	AcAACAAAGTA	CAAaCTGAAA	CTGGAGAGAC	3100																																																																																																																																									
3101	TCTTCGACAA	GAAGAAGACA	GTCGGCAAGT	ATAAGATGCC	GATCAAGAAG	tTGACTCTGC																																																																																																																																								
	TCAAGGAGCA	GCTtGAaAGC	AAaCTCAACT	caCTGAACAA	3200																																																																																																																																									
3201	TCCgAAaCAC	GTaCTGCAgA	ACTTCtcaGT	GTCTTCAAC	AAGAAGAAGG	AaGCCGAGAT																																																																																																																																								
	CGCCGAGACA	GAGAACACTC	TGGAGAACAC	CAAGATTCTt	3300																																																																																																																																									
3301	CTCAAaCACT	ACAAaGGCCT	CGTCAAGTAT	TATAATGGCG	AGTCTTCTCC	TCTGAAGACT																																																																																																																																								
	CTCTCCGAGG	AGAGCATCCA	GACCGAGGAT	AACTACGCCA	3400																																																																																																																																									
3401	GCCTCGAGAA	CTTCAGGTC	CTGTCTAAGC	TCGAAGGCAA	GCTGAAGGAC	AACCTGAACC																																																																																																																																								
	TGGAGAGAA	GAAGCTCAGC	TACCTCTCTA	CGGGACTGCA	3500																																																																																																																																									
3501	TCACCTGATC	GCCGAGCTCA	AGGAAGTCAT	TAAGAACAAAG	AACTACACCG	GCAATAGCCC																																																																																																																																								
	AAGCGAGAA	AATACAGACG	TGAATAACGC	ACTGGAATCT	3600																																																																																																																																									
3601	TatAAGAAGT	TCCCTGCTGA	AGGAACAGAT	GTGCCACTG	TGGTGTCTGA	ATCTGGCTCC																																																																																																																																								
	GACACACTGG	AGCAGTCTCA	ACCTAACAGAAG	CCTGCATCTA	3700																																																																																																																																									
3701	CTCATGTCGG	AGCCGAGTCC	AATACAAATT	CCACATCTCA	GAACGTCGAC	GATGAGGTG																																																																																																																																								
	ATGACGTCT	CATTGTCGCT	ATCTTCGGCG	AGAGCGAGGA	3800																																																																																																																																									
3801	GGACTACGAT	GACCTCGGCC	AGGTGGTCAC	CGGAGAGGCT	GTCACTCCTT	CCGTGATTGA																																																																																																																																								
	TAACATTCTG	TCCAAAATCG	AGAACGAATA	CGAACGTGCTC	3900																																																																																																																																									
3901	TATCTGAAAC	CTCTGGCAGG	CGTCTATAGG	TCTCTCAAGA	AAACAGCTGGA	GAATAACGTG																																																																																																																																								
	ATGACCTCTA	ATGTCACGT	GAAGGACATT	CTGAACAGCC	4000																																																																																																																																									
4001	GCTTTAATAA	GAGAGAAAAT	TTCAAGAACG	TCTTGGAGAG	CGACTTGATT	CCCTATAAAAG																																																																																																																																								
	ACCTGACCTC	CTCTAACTA	ttGTGCAAGG	ACCCATACAA	4100																																																																																																																																									
4101	GTTCCCTCAAT	AAAGAGAAGA	GGGATAAATT	TCTGTCTAGC	TACAACATATA	TCAAGGACTC																																																																																																																																								
	CATCGACACC	GATATCAATT	TCGCTAATGA	TGTGCTGGG	4200																																																																																																																																									
4201	TATTACAAGA	TCCTGAGCGA	AAAATACAAG	TCTGACCTTG	ACTCTATTAA	AAAGTATATC																																																																																																																																								
	AACGATAAGC	AAGGCAGAGA	TGAAAAATAT	CTGCCCTTCC	4300																																																																																																																																									
4301	TGAATAACAT	CGAAACCTTG	TACAAGACAG	TGAACGACAA	AAATCGACCTC	TTCGTaATTC																																																																																																																																								
	ACCTGGAGGC	CAAGGTCCTC	AACTATACTT	ACGAGAAAGAG	4400																																																																																																																																									
4401	CAATGTGAA	GTAAAAATCA	AGGAGCTGAA	CTACCTCAAA	ACAATCCAAG	ACAAGCTGGC																																																																																																																																								
	AGATTTCAAG	AAAATAAAACA	ATTTCGTCGG	AATTGCAAGAC	4500																																																																																																																																									
4501	CTGTCTACCG	ATTATAACCA	CAACAATCTC	CTGACCAAGT	TTCTGTCCAC	TGGCATGGTG																																																																																																																																								
	TTCGAAAACC	TCGCCAAAAC	AGTGCTGAGC	AATCTGCTCG	4600																																																																																																																																									
4601	ACGGCAACCT	GCAGGGCATG	CTGAACATCT	CCCAGCACCA	ATGCGTGAAG	AAACAGTGCC																																																																																																																																								
	CCCAGAAATAG	CGGCTGTTTC	AGGCATCTGG	ACGAGCGCGA	4700																																																																																																																																									
4701	AGAGTGCAAG	TGTCCTCTGA	ACTACAAACA	AGAAGGAGAT	AAAGTGCCTGG	AGAACCCAAA																																																																																																																																								
	CCCTACCTGC	AATGAAAACA	ATGGCGGGTG	TGACGCCGAT	4800																																																																																																																																									
4801	GCTAAATGCA	CCGAGGAAGA	CAGCGGCTCT	AACGGAAAGA	AAATCACATG	CGAGTGTACT																																																																																																																																								
	AAGCCCCACT	CCTATCCACT	CTTcgacggg	atCttCtgct	4900																																																																																																																																									
4901	ccagctctAG	CAACgttACT	ACTTCCGGCA	CTACCCGTCT	TCTATCTGGT	CACACGTGTT																																																																																																																																								
	TCACGTTGAC	AGGTTTGCTT	GGGACGCTAG	TAACCATGGG	5000																																																																																																																																									
5001			CTTGCTGACT			TAA																																																																																																																																								
5013																																																																																																																																														
	70		80		90		100		110		120		130		140		150		160		170		180		190		200		210		220		230		240		250		260		270		280		290		300		310		320		330		340		350		360		370		380		390		400		410		420		430		440		450		460		470		480		490		500		510		520		530		540		550		560		570		580		590		600		610		620		630		640		650		660		670		680		690		700		710		720		730		740		750		760		770	

Figure 32 (Contd.)

	10	20	30	40	50	60	70
		80		90		100	

```

1 atgatgagga aactggccat cctgagcggtg agcagcttcc tgttcggtta ggcctgttt
caggagtacc agtgcgtacgg cagcagcggc aacaccgggg 100
101 tgctgaacga gctgaactac gacaacggcg gcaccaacct gtacaacgag ctggagatga
actactacgg caagcaggag aactggtaca gcctgaagaa 200
201 gaacagccgg tctctggcg agaacgcgca cgccaacaac aacaacggcg acaacggccg
ggaggggcaag gacgaggaca agcggggacgg caacaacgag 300
301 gacaacgaga agctgcggaa gcccacggcac aagaaaactta agcagccgc cgacggcaac
ccccacccca acgccaaccc caacgtggac cccaaacgcca 400
401 atcctaattgt cgaccccaat gccaatccga acgttgatcc caatgcgaat cctaacgcta
accccaatgc caacccaaat gccaatccaa atgcaaatcc 500
501 caacgccaat ccaaacgcaa accctaattgc taatccaaac gctaattccta atgccaatcc
caatgctaac ccaaacgtcg atcctaacgc aaatccgaac 600
601 gctaacccca acgcaaattcc caacgctaac ccgaacgcaa accctaacgc caatccgaat
gccaacccaa acgccaaccc gaacgctaatt ccgaatgtca 700
701 acccgaatgc taatccaaat gcaaacccca aCgcaaaccc caatgcaaac ccAaaTgcca
atcccaacgc caatccctaatt gccaacaaga acaatcagg 800
801 caacggcccg ggcacacaaca tgcccaacga ccccaacccgg aacgtggacg agaacgccaa
cgccaaacgc gccgtgaaga acaacaacaa cgaggagccc 900
901 agcgacaacgc acatcaagga gtacctgaac aagatccaga acagcctgag caccgagtgg
agccccctgca gctgtacatcg cggcaacggc attcagtg 1000
1001 ggatcaagcc cgccagcgcc aacaagccca aggacgagct ggactacgcc aatgacatcg
agaagaagat ctgcaagatg gagaagtgcgac gcaacgtgtt 1100
1101 caacgtggtg aactcctga
1119

```

70		80		90		100		30		40		50		60	
----	--	----	--	----	--	-----	--	----	--	----	--	----	--	----	--

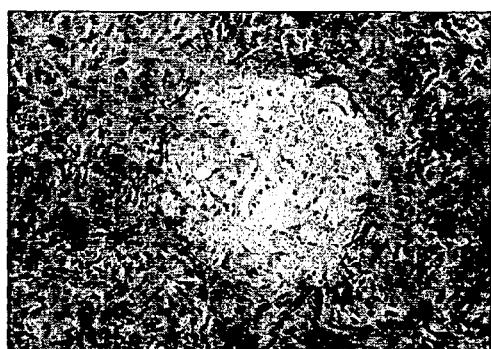
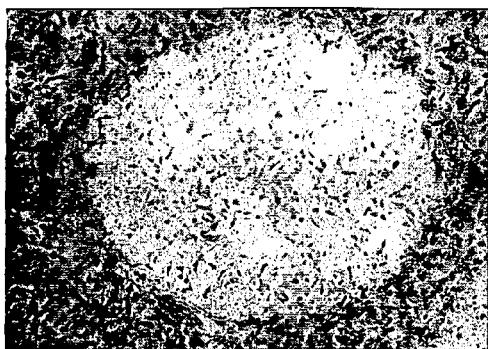
Figure 33

	10	20	30	40	50	60	
70	80	90	100				
<pre> 1 GGTACCGTCA CGCGTCACCG GTGTCATCAT GACCGTGGCC AGGCCCTCTG TGCCTGCCGC CCTGCCCTG CTGGCGAGC TGCCCCGGCT GCTGCTCCTG 100 101 GTGCTGCTGT GCCTGCCCCGCGTGTGGGA TCCGTGATCG AGATCGTGGAA GCGGAGCAAC TACATGGGCA ACCCCTGGAC CGAGTACATG GCCAAGTAGC 200 201 ACATCGAGGA AGTGCACGGC AGCGGCATCC GGGTGGACCT GGGCGAGGAC GCCGAGGTGG CCGGCACCCA GTACAGGCTG CCCAGCGGCAGTGGCCCTG 300 301 GTTCGGCAAG GGCATCATCA TCGAGAACAG CCAGACCACC TTCCCTGACCC CCGTGGCCAC CGAGAACCAAG GACCTGAAGG ACGGGGGCTT CGCCCTTCCC 400 401 CCCACCAAGC CCCTGATGAG CCCCATGACC CTGGACCAGA TGCGGCACCTT CTACAAGGAC AACGAGTACG TGAAGAACCT GGACGAGCTG ACCCTGTGCA 500 501 GCCGGCACGC CGGCAACATG AACCCCGACA ACCGACAAGAA CAGCAACTAC AAGTACCCCG CCGTGTACGA CGACAAGGAT AAGAAGTGCC ACATCCTGTA 600 601 TATCGCCGCC CAGGAAAACA ACGGGCCCCAG GTACTGCAAC AAGGACGAGA GCAAGCGGAA CAGCATGTTG TGCTTCAGAC CCGCCAAGGA CAAGAGCTTC 700 701 CAGAACTACG TGACCTGAG CAAGAACCGTG GTGGACAAC GGGAGAAAGT GTGCCCGGG AAGAACCTGG AAAACGCCAA GTTCGGCCTG TGGGTGACG 800 801 GCAACTGCGA GGACATCCCCCACGTGAACG ACTTCAGCGC CAACGACCTG TTGAGTGCA ACAAGCTGGT GTTCGAGCTG TCCGCCAGCG ACCAGCCCAA 900 901 GCAGTACGAG CAGCACCTGA CCGACTACGA GAAGATCAAA GAGGGCTTCA AGAACAAAGAA CGCCGACATG ATCAAGAGCG CCTTTCTGCC AACTGGCGCC 1000 1001 TTCAAGGCCG ACAGATAACAA GAGCCACCGC AAGGGCTACA ACTGGGGCAA CTACAACAGA AAGACCCAGA AGTGCAGAGAT CTTCAACGTG AAGCCCACCT 1100 1101 GCCTGATCAA CGACAAGTCC TATATGCCCA CCACCGCCCT GAGCCACCCC ATCGAGGTGG AGCACAACCTT CCCTTGCGAC CTGTACAAGG ATGAGATCAA 1200 1201 GAAAGAGATC GAGCGGGAGA GCAAGAGGAT CAAGCTGAAC GACAACGACG ACGAGGGCAA CAAGAAAGATC ATTGCCCGCA GGATCTTCAT CAGCGACGAT 1300 1301 AAGGACAGCC TGAAGTGGCC CTGCGACCC GAGATCGTGT CCCAGAGCAC CTGCAATTTC TTCTGTGCA ATGCGTGGGA AAAGCGGGCC GAAGTGACCA 1400 1401 GCAACAAACGA GGTGGTGGTG AAAGAGGAAT ATAAGGACGA GTACGCCGAC ATCCCCGAGC ACAAGCCCAC CTACGACAAG ATGAGATCA TCATTGCCAG 1500 1501 CTCTGCCGCC GTGGCCGTGC TGGCCACCAT CCTGATGGTG TACCTGTACA AGCGGAAGGG CAACGCCGAG AAGTACGATA AGATGGACCA GCCTCAACGAC 1600 1601 TACGGCAAGA GCACCAAGCCG GAACGACCGAG ATGCTGGACC CCGAGGCCAG CTTCTGGGGC GAGGAAAGA GAGCTAGCCA CACCAACCCCCC GTGCTGATGG 1700 1701 AAAAGCCCTA CTACTGATGA GCGCGCCTGA GCTC 1734 </pre>							
70	80	90	100	30	40	50	

Figure 34

	10	20	30	40	50	60	70
	80	90	100				
1	GGTACCGTCA	CGCGTCACCG	GTGTCATCAT	GACCGTGGCC	AGGCCCTCTG	TGCCTGCCGC	
	CCTGCCCTG	CTGGCGAGC	TGCCCCGGCT	GCTGCTCCTG	100		
101	GTGCTGCTGT	GCCTGCCCGC	CGTGTGGGA	TCCGTGATCG	AGATCGTGGA	GCGGAGCAAC	
	TACATGGGCA	ACCCCTGGAC	CGAGTACATG	GCCAAGTACG	200		
201	ACATCGAGGA	AGTGCACGGC	AGCGGCATCC	GGGTGGACCT	GGGCGAGGAC	GCCGAGGTGG	
	CCGGCACCCA	GTACAGGCTG	CCCAGCGGC	AGTCCCCGT	300		
301	GTTCGGCAAG	GGCATCATCA	TCGAGAACAG	CCAGACCACC	TTCCTGACCC	CCGTGGCCAC	
	CGAGAACCAAG	GACTGAAGG	ACGGCGGCTT	CGCCTTCCCC	400		
401	CCCACCAAGC	CCCTGATGAG	CCCCATGACC	CTGGACCAGA	TGCGGCACCTT	CTACAAGGAC	
	AACGAGTACG	TGAAGAACCT	GGACGAGCTG	ACCCCTGTGCA	500		
501	GCCGGCACGC	CGGCAACATG	AACCCCGACA	ACGACAAGAA	CAGCAACTAC	AAGTACCCCG	
	CCGTGTACGA	CGACAAGGGAT	AAGAAAGTGCC	ACATCCTGTA	600		
601	TATCGCCGCC	CAGGAAAACA	ACGGCCCCAG	GTACTGCAAC	AAGGACGAGA	GCAAGCGGAA	
	CAGCATGTT	TGCTTCAGAC	CCGCCAAGGA	CAAGAGCTTC	700		
701	CAGAACTACG	TGTACCTGAG	CAAGAACGTG	GTGGACAAC	GGGAGAAAGT	GTGCCCCCGG	
	AAGAATCTGG	AAAACGCCAA	GTTCGGCCTG	TGGGTGGACG	800		
801	GCAACTCGCA	GGACATCCCC	CACGTGAACG	AGTTCAGCGC	CAACGACCTG	TTCGAGTGCA	
	ACAAGCTGGT	GTTCGAGCTG	TCCGCCAGCG	ACCAGCCCAA	900		
901	GCAGTACGAG	CAGCACCTGA	CCGACTACGA	GAAGATCAAA	GAGGGCTTCA	AGAACAAAGAA	
	CGCCGACATG	ATCAAGAGCG	CTTTCTGCGC	AACTGGCGCC	1000		
1001	TTCAAGGCCG	ACAGATACAA	GAGCCACGGC	AAGGGCTACA	ACTGGGGCAA	CTACAACAGA	
	AAGACCCAGA	AGTCCGAGAT	CTTCAACGTG	AAGCCCACCT	1100		
1101	GCCTGATCAA	CGACAAGTCC	TATATGCCA	CCACCGCCCT	GAGCCACCCC	ATCGAGGTGG	
	AGCACAACTT	CCCTTGCAGC	CTGTACAAGG	ATGAGATCAA	1200		
1201	GAAAGAGATC	GAGCGGGAGA	GCAAGAGGAT	CAAGCTGAAC	GACAACGACG	ACGAGGGCAA	
	CAAGAAAGATC	ATTGCCCCA	GGATCTTCAT	CAGCGACGAT	1300		
1301	AAGGACAGCC	TGAAGTGCC	CTGCGACCCC	GAGATCGTGT	CCCAGAGCAC	CTGCAATTTC	
	TTCGTGTGCA	AATGCGTGGA	GAAGCGGGGC	GAAGTGACCA	1400		
1401	GCAACAACGA	GGTGGTGGTG	AAAGAGGAAT	ATAAGGACGA	GTACGCCGAC	ATCCCCGAGC	
	ACAAGCCCCAC	CTACGACAAG	ATGTGATGAT	GAGCGCGCCT	1500		
1501						GAGCTC	
1506							
	10	20	30	40	50	60	70
	80	90	100				

Figure 35



rMV3-d-42 3D7

Figure 36

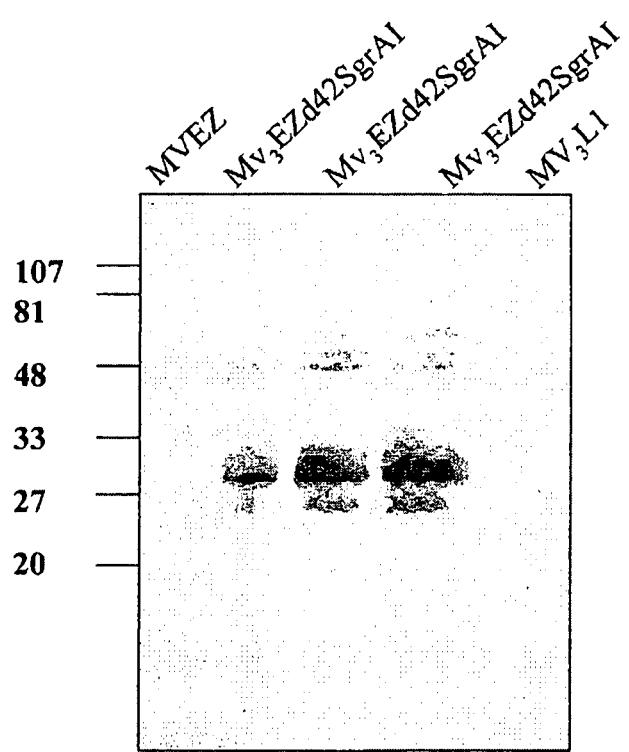


Figure 37

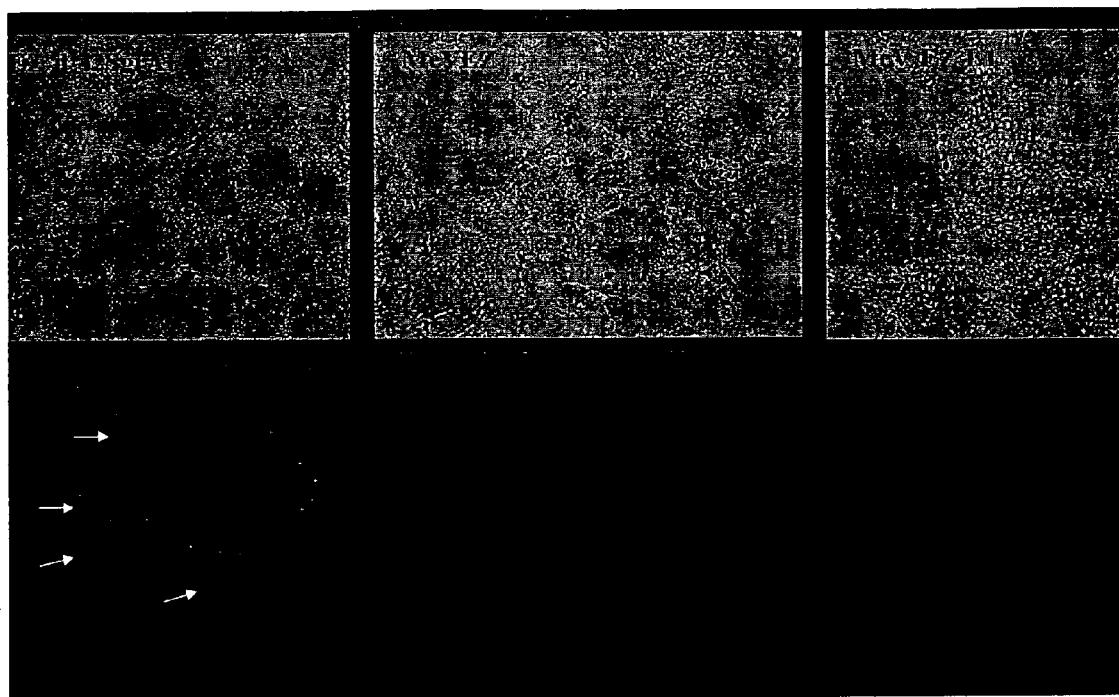


Figure 38

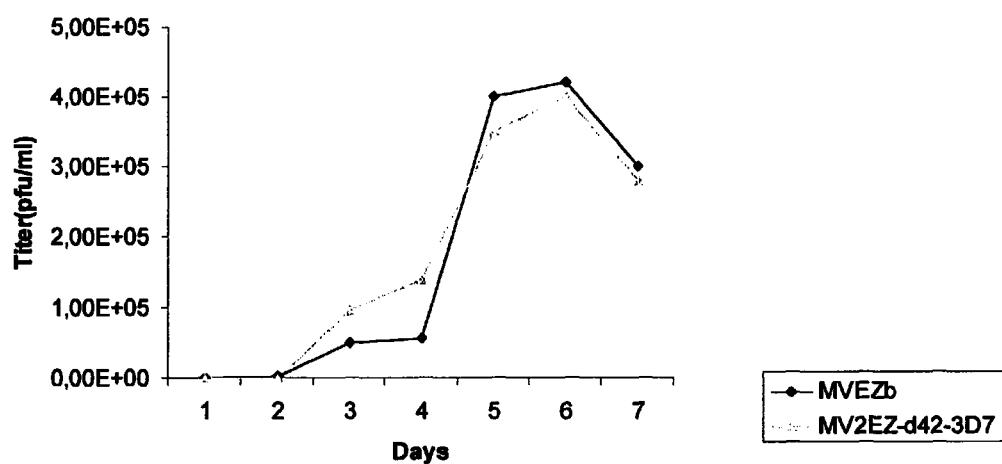


Figure 39

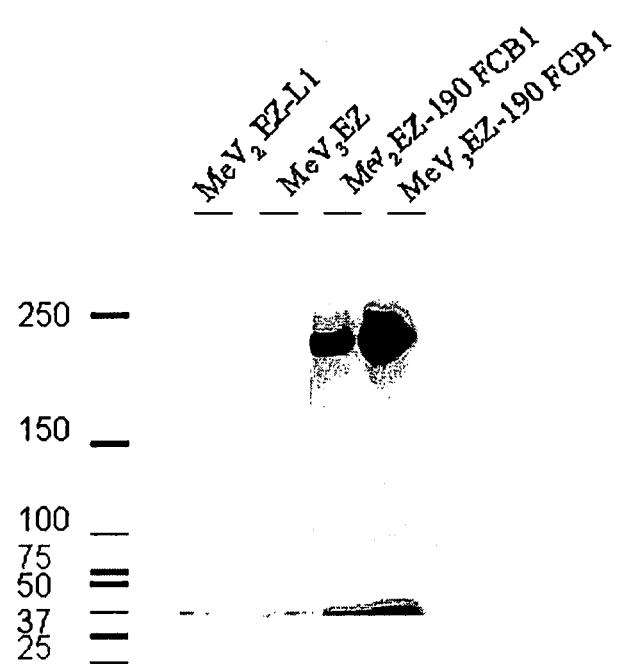


Figure 40

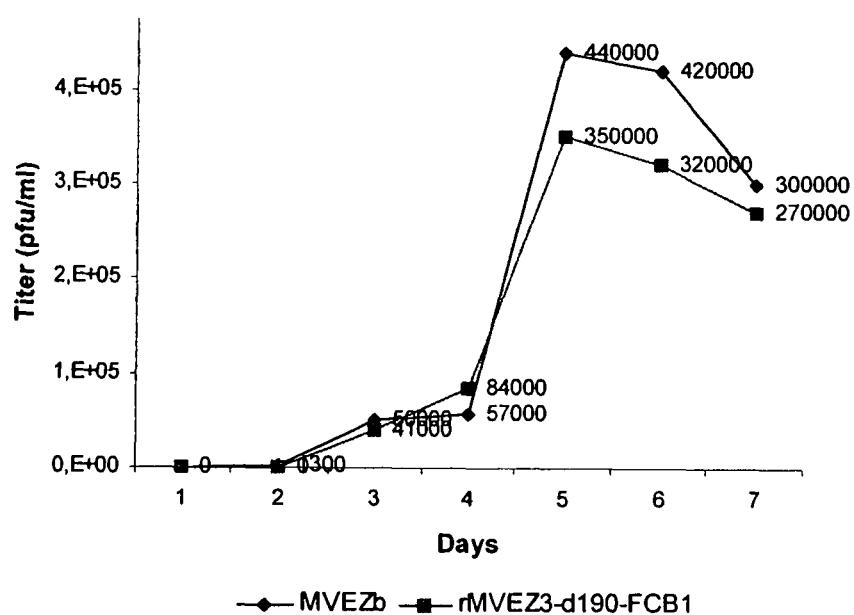
**Figure 41**

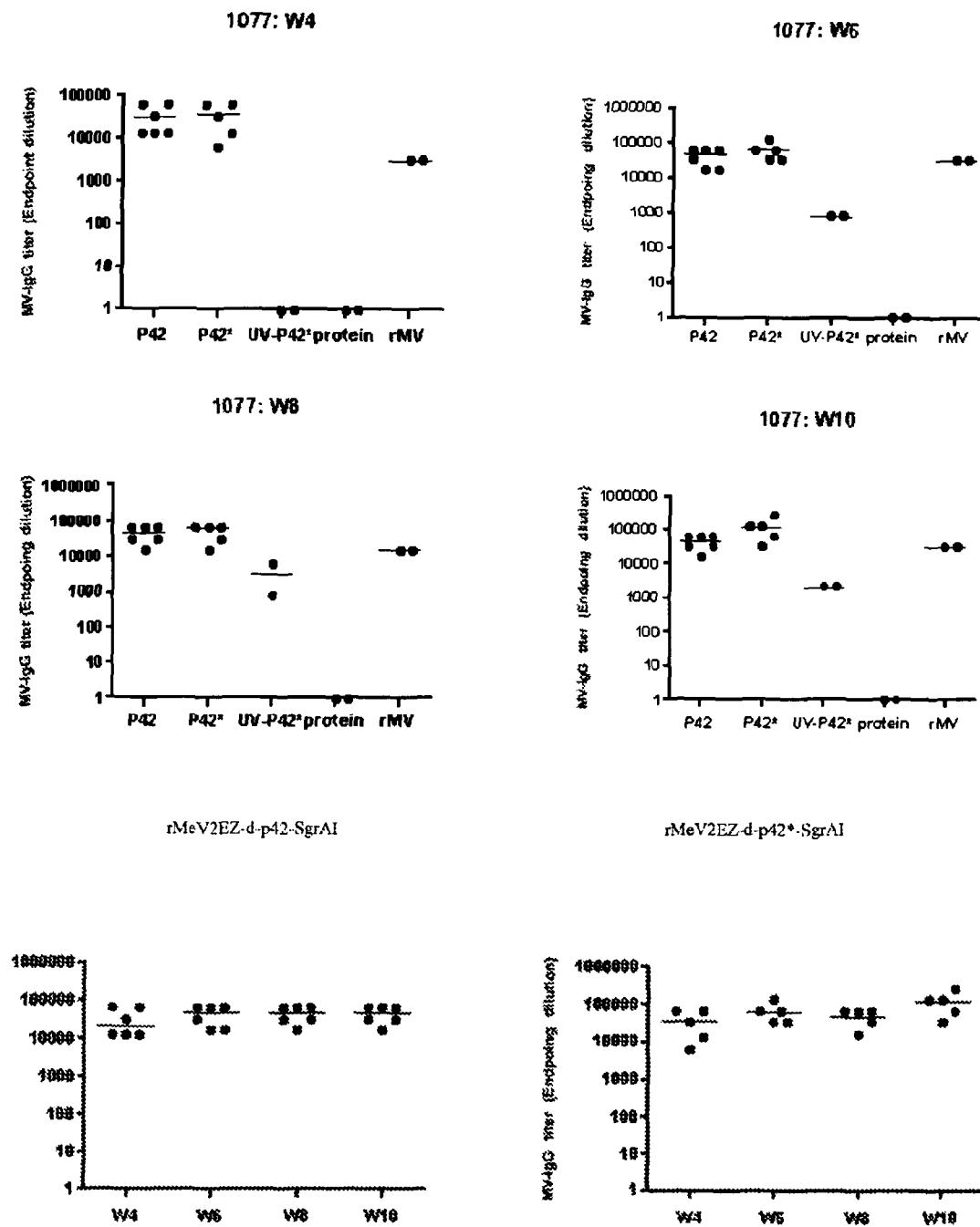
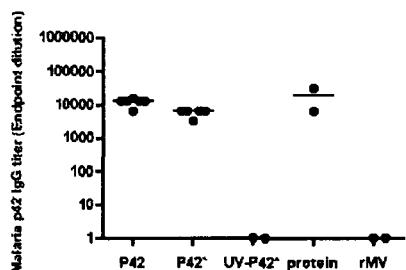
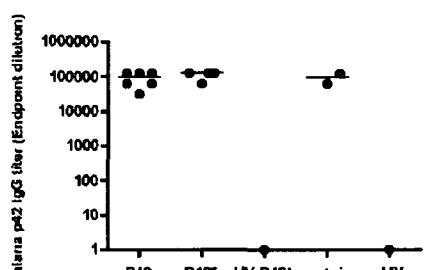
FIGURE 42

FIGURE 43

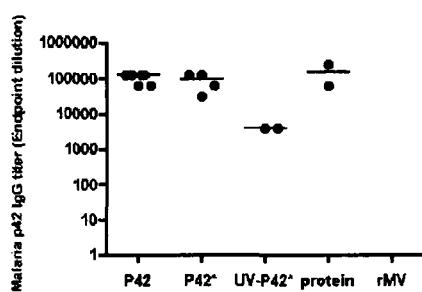
W4



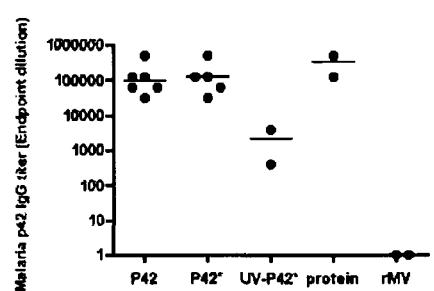
W6



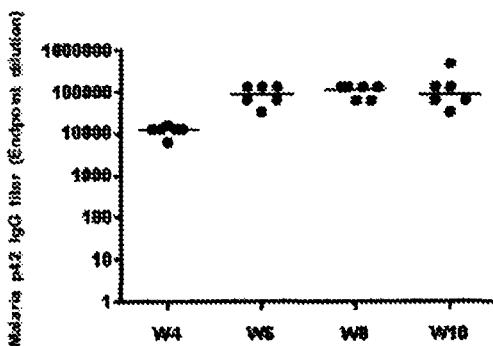
W8



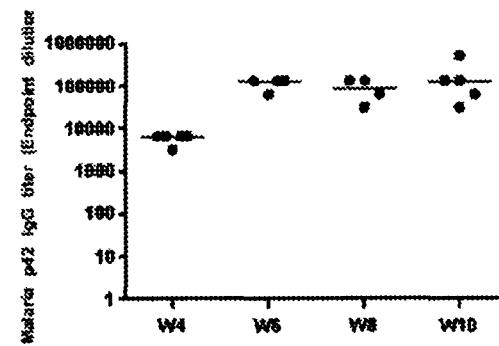
W10



rMeV2EZ-d-p42-SgrAI



rMeV2EZ-d-p42*-SgrAI



COMBINED MEASLES-MALARIA VACCINE

FIELD OF THE INVENTION

The present invention relates to a combined measles-malaria vaccine containing different attenuated recombinant measles-malaria vectors comprising a heterologous nucleic acid encoding several *Plasmodium falciparum* antigens. Preferably, it relates to viral vectors that comprise nucleic acids encoding the circumsporozoite (CS) protein of *P. falciparum*, the merozoite surface protein 1 (MSP-1) of *P. falciparum*, and its derivatives (p-42; p-83-30-38) in its glycosylated and secreted forms, and apical membrane antigen1 (AMA1) of *P. falciparum*, in its anchored or secreted form. The viral vector stems from an attenuated measles virus, based on a strain that is used as a vaccine and is efficient in delivering the gene of interest and that binds to and infects the relevant immune cells efficiently. In a preferred embodiment, the CS, the MSP1 and the AMA1 proteins are generated from the virus such that they will give rise to a potent immune response in mammals, preferably humans; the expression of the proteins is elevated due to human codon optimisation. Furthermore, the invention relates to the use of the recombinant vaccine in the prophylactic treatment of malaria.

BACKGROUND INFORMATION

Measles Virus

The invention relates to a vaccine containing recombinant attenuated measles viruses expressing antigens of *Plasmodium falciparum* (Pf) and to their use for the preparation of recombinant measles-malaria vaccine which will confer immunity against both Measles and Malaria antigens.

Measles virus (MV) is a member of the order Mononegavirales, i.e. viruses with a non-segmented negative-strand RNA genome. The non segmented genome of MV has an antimesage polarity; thus, the genomic RNA is not translated either in vivo or in vitro. Furthermore, it is biologically active only when it is very specifically associated with three viral proteins in the form of a ribonucleoprotein (RNP) complex (see below). Transcription and replication of non-segmented (-) strand RNA viruses and their assembly as virus particles have been reviewed extensively (1). Transcription and replication of measles virus do not involve the nucleus of the infected cells but rather take place in the cytoplasm of infected cells. The genome of the measles virus comprises genes encoding six major structural proteins from the six genes (designated N, P, M, F, H and L) and additionally two non structural proteins derived from the P gene, C and V, involved in counteracting the constitutive immune responses and in regulation of transcription/replication. The gene order is 3' N, P (including C and V), M, F, H, and L 5'. In addition, from the 3'-terminal region a short leader RNA of about 50 nucleotides is transcribed. The cited genes respectively encode the proteins of the ribonucleocapsid (RNP) of the virus, i.e., the nucleoprotein (N), the phosphoprotein (P), and the large polymerase/replicase protein (L), which very tightly associate with the genome RNA, forming the RNP. The other genes encode the proteins of the viral envelope including the hemagglutinin (H), the fusion (F) and the matrix (M) proteins. The transcription of the MV genes follows a decreasing gradient: when the polymerase operates on the genomic template it synthesizes more RNA made from upstream genes than from downstream genes. In this discontinuous transcription mode the mRNAs are capped and polyadenylated. Conversely, in the replication mode, the L protein produces full

length antigenomic and genomic RNA which are immediately covered with N, P and L proteins to form infectious progeny RNPs.

The measles virus has been isolated in 1954: Enders and Peebles inoculated primary human kidney cells with the blood of David Edmonston, a child affected by measles, and the resulting Edmonston strain of MV (2) was subsequently adapted to growth in a variety of cell lines. Adaptation to chicken embryos, chick embryo fibroblasts (CEF), and/or dog kidney cells and human diploid cells produced the attenuated Edmonston A and B (3), Zagreb (EZ) and AIK-C seeds. Edmonston B was licensed in 1963 as the first MV vaccine. Further passages of Edmonston A and B on CEF produced the more attenuated Schwarz and Moraten viruses (3) whose sequences have recently been shown to be identical (4; 5). Because Edmonston B vaccine was reactogenic, it was abandoned in 1975 and replaced by the Schwarz/Moraten vaccine. Several other vaccine strains are also used: AIK-C, Schwarz F88, CAM70, TD97 in Japan, Leningrad-16 in Russia, and 10 Edmonston Zagreb. The CAM70 and TD97 Chinese strains were not derived from Edmonston. Schwarz/Moraten and AIK-C vaccines are produced on CEF. Zagreb vaccine is produced on human diploid cells (WI-38). Today, the Schwarz/Moraten, AIK-C and EZ vaccines are commonly 15 used (6), but in principle, any one of these attenuated vaccine strains, which are all of the one unique MV serotype, proven to be safe and to induce long-lasting immune responses, can be used for the purposes of the invention.

MV vaccines induce life-long immunity after a single or 20 two low-dose injections. Protection against measles is mediated both by antibodies and by CD4 and CD8 T cells. Persistence of MV-specific antibodies and CD8 cells has been shown for as long as 25 years after vaccination (7).

MV vaccine is easy to produce on a large scale in most 25 countries and can be distributed at low cost. Because the attenuation of MV genome results from an advantageous combination of numerous mutations, the vaccine is very stable and reversion to pathogenicity has never been observed (6).

Regarding safety, MV replicates exclusively in the cytoplasm, ruling out the possibility of integration into host DNA. These characteristics make live attenuated MV vaccine an attractive candidate to be used as a multivalent vaccination vector. Such a vaccine may prove as efficient in eliciting 35 long-lasting immune protection against other pathogenic agents as against the vector virus itself.

Martin Billeter and colleagues cloned cDNA corresponding 40 to the antigenome of Edmonston MV, and established an original and efficient reverse genetics procedure to rescue the virus (8), as described in International Patent Application WO 97/06270. The recombinant measles virus is recovered from the helper cell line 293-3-46, stably transfected and expressing MV N and P proteins as well as bacteriophage T7 RNA polymerase. For rescue of any variant or recombinant MV the 45 helper cell line is then transiently transfected with an expression plasmid encoding L protein, and most importantly with any antigenomic plasmid appropriately constructed to yield any mutated or recombinant antigenomic RNA compatible to give rise to progeny MV. The transient transfection step leads 50 first to the transcription, preferably by the resident T7 RNA polymerase. The resulting antigenomic RNA is immediately (in statu nascendi) covered by the viral N, P and L proteins, to yield antigenomic RNP from which genomic RNP is produced. Second, the genomic RNP is transcribed by the 55 attached L, to yield all viral mRNAs and the respective proteins. Finally, both genomic and antigenomic RNPs are amplified by replication.

In a slight variation of this procedure, rather than using stably transfected 293-3-46 helper cells, commercially available 293T cells have been transiently transfected, using simultaneously all 5 plasmids detailed in the original patent description, those encoding N, P and T7 polymerase (previously used to create the helper cell line) as well as the plasmid encoding L and the antigenomic plasmid. Note that in the "fully transient transfection" procedure it is possible to use also variant expression plasmids and to avoid the use of T7 RNA polymerase altogether, utilizing instead the resident RNA polymerase H to express also the L protein and the antigenome (9).

To rescue individual recombinant MVs the antigenomic plasmids utilized comprise the cDNA encoding the full length antigenomic (+)RNA of the measles virus recombined with nucleotide sequences encoding the heterologous antigen of interest (heterologous nucleotide sequence), flanked by MV-specific transcription start and termination sequences, thus forming additional transcription units (ATUs). This MV Edmonston strain vector has been developed by the original MV rescue inventors for the expression of foreign genes (10), demonstrating its large capacity of insertion (as much as 5 kb) and the high stability in the expression of transgenes (11; 12), such as Hepatitis B virus surface antigen, simian or human immunodeficiency viruses (SIV or HIV), mumps virus, and human IL-12. In particular, early on, recombinant measles virus expressing Hepatitis B virus surface and core antigens either individually or in combination have been produced and shown to induce humoral immune responses in genetically modified mice.

From the observation that the properties of the measles virus and especially its ability to elicit high titers of neutralizing antibodies in vivo and its property to be a potent inducer of long lasting cellular immune response, the inventors have proposed that it may be a good candidate for the production of recombinant viruses expressing antigens from *P. falciparum*, to induce neutralizing antibodies against said Malaria parasite which preferably could be suitable to achieve at least some degree of protection in animals and more preferably in human hosts.

Especially, MV strains and in particular vaccine strains have been elected in the present invention as candidate vectors to induce immunity against both measles virus and *P. falciparum* parasite whose constituent is expressed in the designed recombinant MV, in exposed infant populations because they are having no MV immunity.

Adult populations, even already MV immunized individuals, may however also benefit from MV recombinant immunization because re-administering MV virus under the recombinant form of the present invention results in a boost of anti-MV antibodies (13).

The invention relates in particular to the preparation of recombinant measles viruses bearing heterologous genes from *P. falciparum* parasites.

The advantageous immunological properties of the recombinant measles viruses according to the invention can be shown in an animal model which is chosen among animals susceptible to measles viruses, and wherein the humoral and/or cellular immune response against the heterologous antigen and/or against the measles virus is determined. Among such animals suitable to be used as model for the characterization of the immune response, the skilled person can especially use transgenic mice expressing CD46, one of the specific receptors for MV. The most promising recombinants can then be tested in monkeys.

The recombinant measles virus nucleotide sequence must comprise a total number of nucleotides which is a multiple of

six. Adherence to this so-called "rule of six" is an absolute requirement not only for MV, but for all viruses belonging to the subfamily Paramyxovirinae. Apparently, the N protein molecules, each of which contacts six nucleotides, must cover the genomic and antigenomic RNAs precisely from the 5' to the 3' end.

It is of note that the location of the ATUs can vary along the antigenomic cDNA. Thus, taking advantage of the natural expression gradient of the mRNAs of MV mentioned above, the level of expression of inserted ATUs can be varied to appropriate levels. Preferred locations of ATUs are upstream of the L-gene, upstream from the M gene and upstream of the N gene, resulting in low, medium and strong expression, respectively, of heterologous proteins.

Malaria Parasite

Malaria currently represents one of the most prevalent infectious diseases in the world, especially in tropical and subtropical areas. Per year, malaria infections lead to severe illnesses in hundreds of million individuals worldwide, killing between 1 and 3 million, primarily young infants in developing and emerging countries. The widespread occurrence and elevated incidence of malaria are a consequence of the widespread ban of DDT and the increasing numbers of drug-resistant parasites as well as insecticide-resistant parasite vectors. Other factors include environmental and climatic changes, civil disturbances, and increased mobility of populations.

Malaria is caused by the mosquito-borne hematoplasmodial parasites belonging to the genus *Plasmodium* from the phylum Apicomplexa. Four species of *Plasmodium* genus infect humans: *P. malariae*, responsible for Malaria quartana, *P. vivax* and *P. ovale*, both of which cause Malaria tertiana, and *P. falciparum*, the pathogen of Malaria tropica and responsible for almost all fatal infections. Many others cause disease in animals, such as *P. yoelii* and *P. berghei* in mice.

Malaria parasites have a life cycle consisting of several stages. Each stage is able to induce specific immune responses directed against the corresponding occurring stage-specific antigens. Malaria parasites are transmitted to man by several species of female *Anopheles* mosquitoes. Infected mosquitoes inject the "sporozoite" form of the malaria parasite into the mammalian bloodstream. Sporozoites remain for a few minutes in the circulation before invading hepatocytes. At this stage, the parasite is located in the extra-cellular environment and is exposed to antibody attack, mainly directed to the "circumsporozoite" (CS) protein, a major component of the sporozoite surface. Once in the liver, the parasites replicate and develop into so-called "schizonts." These schizonts occur in a ratio of up to 20,000 per infected cell. During this intra-cellular stage of the parasite, main players of the host immune response are T-lymphocytes, especially CD8+ T-lymphocytes. After about one week of liver infection, thousands of so-called "merozoites" are released into the bloodstream. Apical membrane antigen 1 (AMA1) and merozoite surface protein 1 (MSP1) are both present on merozoites that emerge from infected liver cells: they are essential components of the asexual blood-stage merozoite, responsible for invasion of erythrocytes. Once they enter red blood cells, they become targets of antibody-mediated immune response and T-cell secreted cytokines. After invading erythrocytes, the merozoites undergo several stages of replication, giving rise to so-called "trophozoites" and to schizonts and merozoites, which can infect new red blood cells. A limited amount of trophozoites may evolve into "gametocytes," which constitute the parasite's sexual stage. When susceptible mosquitoes ingest erythrocytes, gametocytes are released from the erythrocytes, resulting in several male gametocytes and one female

gametocyte. The fertilization of these gametes leads to zygote formation and subsequent transformation into ookinets, then into oocysts, and finally into salivary gland sporozoites. Targeting antibodies against gametocyte stage-specific surface antigens can block this cycle within the mosquito mid gut. Such antibodies will not protect the mammalian host but will reduce malaria transmission by decreasing the number of infected mosquitoes and their parasite load.

The MSP-1 is synthesised as 190-200 kDa (d-190) precursor which is proteolytically processed into fragments of 83, 30, 38 and 42 kDa (d-42) during schizogony (14). At the time of erythrocytic invasion the 42-kDa is further cleaved to yield a 33 kDa fragment which is shed with the rest of the complex, and a 19 kDa fragment, which contains two epidermal growth factor (EGF)-like domains, that remains associated with the merozoite membrane during invasion. This secondary cleavage is a pre-requisite for successfully erythrocyte invasion (15).

MSP-1 is an essentially dimorphic protein exhibiting high conservation within the dimorphic alleles characterised by the K1 and MAD20 prototypes.

AMA-1 (16) is a structurally conserved type I integral membrane protein, comprising 622 aa in *P. falciparum* (PfAMA-1), organised in a cytosolic region (50 aa), a transmembrane region, and an ectodomain, which folds as an N-terminal pro-sequence and three domains (DI, DII, DIII). Expression of the protein is maximal in late schizogony: the precursor of AMA-1 (83 kDa) is processed proteolytically, to cleave away the pro-sequence, converting the protein into a 66 kDa form, which allows the merozoite relocalisation. Antibodies recognise mainly DI and DII, and appear to react equally well with several allelic variants. Antibody responses to DIII are generally low, levels increasing in adults (17, 18).

PfAMA-1 contains 64 polymorphic positions (9 in the pro-sequence, 52 in the ectodomain, 3 in the cytosolic region), most of them are dimorphic, which are important epitopes for host immune responses. To develop PfAMA-1-based vaccines it should be important to cover the polymorphisms: Diversity Covering (DiCo 1, 2 and 3) PfAMA-1 are artificial sequences representing, to the greatest extent possible, the naturally occurring polymorphism of the PfAMA1 ectodomain. It has been shown that they induce immune responses which are functional against a range of parasites carrying diverse PfAMA1 alleles. This approach may offer a means by which vaccines targeting PfAMA1 can be produced such that a strong and a functional protection against the broad range of naturally occurring PfAMA1 alleles can be induced. (19).

The CS protein (CSP) has about 420 aa and a molecular weight of 58 kDa. It represents the major surface protein of sporozoites: its function is fundamental for the maturation of sporozoites from oocysts and for the invasion of hepatocytes, which is mediated from a conserved motif of positively charged aminoacids. CSP is organised into two non-repetitive regions at 5' and 3' ends, and a variable species-specific central region, consisting of multiple repeats of four-residues-long motifs, which represents the main epitope within the CSP. Since CSP continues to be detectable for at least the first 3 days of schizogony, it is considered an attractive vaccine target for both antibody-mediated immuno response, directed against extracellular sporozoites, and cell-mediated immuno responses, directed against schizonts (20).

Current approaches to malaria vaccine development can be classified according to the different stages in which the parasite can exist, as described above.

Three types of possible vaccines can be distinguished: i) pre-erythrocytic vaccines, which are directed against sporozoites and/or schizont-infected cells. These types of vaccines are primarily CS-based, and should ideally confer sterile immunity, mediated by humoral and cellular immune responses, preventing malaria infection; ii) asexual blood-stage vaccines, which are directed against merozoites-infected cells: MSP1 and AMA1 are leading malaria vaccine candidates, designed to minimize clinical severity. These vaccines should reduce morbidity and mortality and are meant to prevent the parasite from entering and/or developing in the erythrocytes; iii) transmission-blocking vaccines, which are designed to hamper the parasite development in the mosquito host. This type of vaccine should favour the reduction of population-wide malaria infection rates. Next to these vaccines, the feasibility of developing malaria vaccines that target multiple stages of the parasite life cycle is being pursued in so-called multi-component and/or multi-stage vaccines.

Today's global malaria vaccine portfolio looks promising with 47 new vaccine candidates, 31 in preclinical development, narrowing down to 16 in clinical trials. One of these, the RTS,S vaccine, being developed by GSK Biologicals and PATH-MVI, should enter final phase III clinical trials in 2008 (21). Other interesting vaccine candidates are those based on live recombinant viruses used as vector, such as Modified Vaccinia Ankara (MVA), as described in International Patent Application US2006127413, poxvirus (U.S. Pat. No. 6,214,353, AU7060294, AU1668197, WO9428930, and U.S. Pat. No. 5,756,101), adenovirus (US2007071726, US2005265974, US2007088156 and CA2507915), cold-adapted attenuated influenza virus, or based on yeasts, such as *Pichia pastoris* and *Saccharomyces* spp., or on bacterial expression systems, such as *Salmonella* spp. (U.S. Pat. No. 5,112,749) and *Escherichia coli* (EB0191748) (22).

Currently, no commercially available vaccine against malaria is available, although the development of vaccines against malaria has already been initiated more than 30 years ago. Many factors make malaria vaccine development difficult and challenging. First, the size and genetic complexity of the parasite mean that each infection presents thousands of antigens to the human immune system. Understanding which of these can be a useful target for vaccine development has been complicated, and to date at least 40 different promising antigens have been identified. Second, the parasite changes through several life stages even while in the human host, presenting, at each stage of the life cycle, a different subset of molecules to the immune system. Third, the parasite has evolved a series of strategies that allow it to confuse, hide, and misdirect the human immune system. Finally, it is possible to have multiple malaria infections of not only different species but also of different strains at the same time.

Hence the present invention fulfills the long felt need of prior art by providing combined measles-malaria vaccine containing different attenuated recombinant measles-malaria vectors comprising a heterologous nucleic acid encoding several *Plasmodium falciparum* antigens.

SUMMARY OF THE INVENTION

In one embodiment of the present invention provides a combined measles-malaria vaccine comprises a recombinant measles vaccine virus which express malaria antigens capable of eliciting immune response and protection both against measles and malaria.

In another embodiment, the present invention provides the recombinant measles vaccine virus having nucleotide sequence which expresses MSP1 malaria antigen. In preferred embodiment, recombinant measles vaccine virus having nucleotide sequence which expresses malaria antigen

d190 or d83-30-38 or d42 in both anchored and secreted forms from 3D7 strain and the FCB1 strain.

In yet another embodiment, the present invention provides the recombinant measles vaccine virus having nucleotide sequence which expresses Diversity Covering (DiCo) AMA1 malaria antigen.

In yet another embodiment, the present invention provides the recombinant measles vaccine virus having nucleotide sequence which expresses CS malaria antigen.

DESCRIPTION OF THE FIGURES

FIG. 1: Schematic representation of the antigenomic DNA p(+)-MV-EZ of measles virus. p(+)-MV-EZ is a plasmid derived from pBluescript containing the complete sequence of the measles virus (Edmiston Zagreb), under the control of the T7 RNA polymerase promoter (T7), containing three ATU respectively in position 1 (before the N gene of the measles virus), 2 (between the P and the M genes of the measles virus) and 3 (between the H and the L genes of the measles virus), and exactly terminated by the hepatitis delta ribozyme and T7 RNA polymerase terminator (δ T7t). The size of the plasmid is 18941 bp.

FIG. 2: Representation of the MSP-1 synthetic gene (d-190) from 3D7 strain. The coding nucleotides on the flanking regions of the d-190 gene fragments (d-83-30-38 and d-42) and the corresponding aminoacids are shown. Unique restriction sites added for cloning procedures are in colours; SP: signal peptide; GPI: glycosyl-phosphatidil-inositol sequence coded for membrane-anchored region.

FIG. 3: Representation of the MSP-1 synthetic gene (d-190) from FCB1 strain. The coding nucleotides on the flanking regions of the d-190 gene fragments (d-83-30-38 and d-42) and, the corresponding aminoacids are shown. Unique restriction sites added for cloning procedures are in colours; SP: signal peptide; GPI: glycosyl-phosphatidil-inositol sequence coded for membrane-anchored region. SP and GPI regions are from 3D7 strain.

FIG. 4: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-d190-3D7. It is a plasmid derived from p(+)-MV-EZ containing d-190 malaria gene (3D7 strain), 5253 bp, coding the GPI-anchored form of the protein, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 24323 bp.

FIG. 5: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-d190*-3D7. It is a plasmid derived from p(+)-MV-EZ containing d-190* malaria gene (3D7 strain), 5160 bp, coding the secreted form of the protein, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 24227 bp.

FIG. 6: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₃-EZ-d190-3D7 or p(+)-MV₃-EZ-d190*-3D7. It is a plasmid derived from p(+)-MV-EZ containing the d-190 malaria gene (3D7 strain), 5253 bp, coding the GPI-anchored form of the protein, or the d-190* malaria gene (3D7 strain), 5160 bp, coding the secreted form of the protein, cloned in position three of the measles genome by SgrAI-BssHII digestion. The recombinant plasmid p(+)-MV₃-EZ-d190 is 24323 bp, and p(+)-MV₃-EZ-d190* is 24227 bp.

FIG. 7: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-d83-30-8-3D7. It is a plasmid derived from p(+)-MV-EZ containing d-83-30-38 malaria gene (3D7 strain), 4122 bp, coding the GPI-anchored

form of the protein, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 23195 bp.

FIG. 8: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-d83-30-38*-3D7. It is a plasmid derived from p(+)-MV-EZ containing d-83-30-38* malaria gene (3D7 strain), 4029 bp, coding the secreted form of the protein, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 23105 bp.

FIG. 9: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₃-EZ-d83-30-38-3D7 or p(+)-MV₃-EZ-d83-30-38*-3D7. It is a plasmid derived from p(+)-MV-EZ containing d-83-30-38 malaria gene (3D7 strain), 4122 bp, coding the GPI-anchored form of the protein, or the d-83-30-38* gene (3D7 strain), 4029 bp, coding the secreted form of the protein, cloned in position three of the measles genome by SgrAI-BssHII digestion. The recombinant plasmid p(+)-MV₃-EZ-d83-30-38 is 23195 bp, p(+)-MV₃-EZ-d83-30-38* is 23105 bp.

FIG. 10: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-d42-3D7. It is a plasmid derived from p(+)-MV-EZ containing d-42 malaria gene (3D7 strain), 1347 bp, coding the GPI-anchored form of the protein, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 20417 bp.

FIG. 11: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-d42*-3D7. It is a plasmid derived from p(+)-MV-EZ containing d-42* malaria gene (3D7 strain), 1254 bp, coding the secreted form of the protein, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 20345 bp.

FIG. 12: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₃-EZ-d42-3D7 or p(+)-MV₃-EZ-d42*-3D7. It is a plasmid derived from p(+)-MV-EZ containing d-42 malaria gene (3D7 strain), 1347 bp, coding the GPI-anchored form of the protein, or the d-42* malaria gene (3D7 strain), 1254 bp, coding the secreted form of the protein, cloned in position three of the measles genome by SgrAI-BssHII digestion. The recombinant p(+)-MV₃-EZ-d42 is 20417 bp, the p(+)-MV₃-EZ-d42* is 20345 bp.

FIG. 13: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-d190-FCB1. It is a plasmid derived from p(+)-MV-EZ containing d-190 malaria gene (FCB1 strain), 5013 bp, coding the GPI-anchored form of the protein, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 24083 bp.

FIG. 14: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₃-EZ-d190-FCB1. It is a plasmid derived from p(+)-MV-EZ containing the d-190 malaria gene (FCB1 strain), 5013 bp, coding the GPI-anchored form of the protein, cloned in position three of the measles genome by SgrAI-BssHII digestion. The recombinant plasmid p(+)-MV₃-EZ-d190 is 24083 bp.

FIG. 15: Representation of the CS synthetic gene. The coding nucleotides on the flanking regions of the CS gene and the corresponding aminoacids are shown. Unique restriction sites added for cloning procedures are in colours.

FIG. 16: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-CS. It is a plasmid derived from p(+)-MV-EZ containing CS gene, 1119 bp, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 20219 bp.

FIG. 17: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₃-EZ-CS. It is a plasmid derived from p(+)-MV-EZ containing CS gene, 1119 bp, cloned in position three of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 20219 bp.

FIG. 18: Representation of the DiCo-1 complete synthetic gene. The coding nucleotides on the flanking regions of the DiCo1 complete gene domains (ecto and trans-cyto) and the corresponding aminoacids are shown. Unique restriction sites added for cloning procedures are in colours; SP: signal peptide human codon optimised.

FIG. 19: Representation of the DiCo-1 ecto synthetic gene. The coding nucleotides on the flanking regions of the DiCo1 ecto domain and the corresponding aminoacids are shown. Unique restriction sites added for cloning procedures are in colours; SP: signal peptide (human codon optimised).

FIG. 20: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-DiCo1-complete. It is a plasmid derived from p(+)-MV-EZ containing DiCo1 complete gene, 1689 bp, coding the transmembrane form of the protein, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 20753 bp.

FIG. 21: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₃-EZ-DiCo1-complete. It is a plasmid derived from p(+)-MV-EZ containing DiCo1 complete gene, 1689 bp, coding the transmembrane form of the protein, cloned in position three of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 20753 bp.

FIG. 22: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₂-EZ-DiCo1-ecto. It is a plasmid derived from p(+)-MV-EZ containing DiCo1 ecto gene, 1458 bp, coding the secreted form of the protein, cloned in position two of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 20525 bp.

FIG. 23: Schematic representation of the recombinant measles-malaria plasmid, p(+)-MV₃-EZ-DiCo1-ecto. It is a plasmid derived from p(+)-MV-EZ containing DiCo1 ecto gene, 1458 bp, coding the secreted form of the protein, cloned in position three of the measles genome by SgrAI-BssHII digestion. The size of the recombinant plasmid is 20525 bp.

FIG. 24: Complete nucleotide sequence of p(+)-MV₂-EZ-GFP. The sequence can be described as follows with reference to the position of the nucleotides:

592-608 T7 promoter
609-17354 MV Edmoston Zagreb antigenome
4049-4054 MluI restriction site
4060-4067 SgrAI restriction site
4079-4084 BssHII restriction site
4085-4801 Green Fluorescent Protein (GFP) ORF
4805-4810 BssHII restriction site
4817-4822 AatII restriction site

17355-17580 HDV ribozyme and T7 terminator

FIG. 25: Complete nucleotide sequence of p(+)-MV₃-EZ-GFP. The sequence can be described as follows with reference to the position of the nucleotides:

592-608 T7 promoter
609-17359 MV Edmoston Zagreb antigenome
9851-9856 MluI restriction site
9862-9869 SgrAI restriction site
9886-9891 BssHII restriction site
9892-10608 Green Fluorescent Protein (GFP) ORF
10612-10617 BssHII restriction site
10624-10629 AatII restriction site
17360-17585 HDV ribozyme and T7 terminator

FIG. 26: AN101TE: this is the MSP1 d-190 3D7 sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-99 d-1903D7 signal peptide
100-105 BamHI restriction site
4014-4020 BstEII restriction site
5152-5157 AclI restriction site
5158-5250 GPI sequence
5251-5253 STOP codon

FIG. 27: AN102TE: this is the MSP1 d-190* 3D7 sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-99 d-190*3D7 signal peptide
100-105 BamHI restriction site
4014-4020 BstEII restriction site
5152-5157 AclI restriction site
5158-5160 STOP codon

FIG. 28: AN103TE: this is the MSP1 d-83-30-38 3D7 sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-99 d-83-30-38 3D7 signal peptide
100-105 BamHI restriction site
4014-4020 BstEII restriction site
4021-4026 AclI restriction site
4027-4119 GPI sequence
4120-4122 STOP codon

FIG. 29: AN104TE: this is the MSP1 d-83-30-38* 3D7 sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-99 d-83-30-38* 3D7 signal peptide
100-105 BamHI restriction site
4014-4020 BstEII restriction site
4027-4029 STOP codon

FIG. 30: AN105TE: this is the MSP1 d-42 3D7 sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-99 d-42 3D7 signal peptide
100-105 BamHI restriction site
108-114 BstEII restriction site
1246-1251 AclI restriction sites
1252-1344 GPI sequence
1345-1347 STOP codon

FIG. 31: AN106TE: this is the MSP1 d-42* 3D7 sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-99 d-42* 3D7 signal peptide
100-105 BamHI restriction site
108-114 BstEII restriction site
1246-1251 AclI restriction sites
1252-1254 STOP codon

FIG. 32: AN107TE: this is the MSP1 d-190 FCB1 sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-99 d-190 FCB1 signal peptide
100-105 BamHI restriction site
146-151 HindIII restriction site
3825-3831 BstEII restriction site

11

4912-4917 AcII restriction sites
4918-5010 GPI sequence
5011-5013 STOP codon

FIG. 33: AN108TE: this is the CS sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-1116 CS sequence
1117-1119 STOP codon

FIG. 34: AN109TE: this is the DiCo 1 complete sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-99 DiCo1 complete signal peptide
100-105 BamHI restriction site
106-1686 DiCo 1 complete sequence ORF
1687-1689 STOP codon

FIG. 35: AN110TE: this is the DiCo 1 ecto sequence ORF cloned by the inventors. The sequence can be described as follows with reference to the position of the nucleotides:

1-3 Start codon
4-99 DiCo1 ecto signal peptide
100-105 BamHI restriction site
106-1455 DiCo 1 ecto sequence ORF
1456-1458 STOP codon

FIG. 36: Comparable cytopathic effects produced on Vero cells after infection with the recombinant Measles-p-42 Malaria virus MV virus vaccine.

FIG. 37: Expression of the d-42 3D7 transgene inserted into position three of the Measles vector (MV₃EZ-d-42 SgrAI). Cell lysates from passage 1, 5 and 10 analysed by Western Blot against empty Measles vector (MVEZ) and a negative control (MV₃L1, a recombinant MV-Papilloma virus).

FIG. 38: Expression of the d-42 3D7 transgene inserted into position three of the Measles vector (MV₃EZ-d-42 SgrAI) analysed by immunofluorescence, compared with empty Measles vector (MVEZ) and a negative control (MV2EZL1, a recombinant MV-Papilloma virus). Arrows point to the same syncytia as they looked using an optical microscope before and after immunostaining.

FIG. 39: Growth kinetics curve of the recombinant Measles-p-42 Malaria virus compared with that of the MV virus vaccine.

FIG. 40: Expression of the d-190 FCB1 transgene inserted into position two and three of the Measles vector (MV_{2,3}EZ-d-190 SgrAI FCB1). Cell lysates analysed by Western Blot against empty Measles vector (MVEZ) and a negative control (MV2EZL1, a recombinant MV-Papilloma virus).

FIG. 41: Growth kinetics curve of the recombinant Measles-p-190-FCB1 Malaria virus compared with that of the MV virus vaccine.

FIG. 42: Shows humoral immune responses against Measles.

FIG. 43: Shows humoral immune responses against Malaria p42.

DETAILED DESCRIPTION OF THE INVENTION

The object of the invention is the production of a combined measles-malaria vaccine from a recombinant Measles vectors capable of containing stably integrated DNA sequences which code for CS, MSP-1 or partial sections of it and AMA-1 or partial sections, in the secreted or surface anchored forms, of *P. falciparum*.

The invention shall also include the rescue of recombinant MV-Malaria viruses which are capable of infection, replica-

12

tion and expression of PfCS, PfMSP-1 and PfAMA-1 antigens in susceptible transgenic mice, monkeys and human host.

Furthermore, the invention intends to include the construction of multivalent recombinant measles-malaria vectors, in which two different antigens are simultaneously cloned and expressed in the same vector, conferring immunity against both of them.

Moreover, the invention relates to the combination of three 10 different recombinant measles-malaria viruses, each carrying a different gene and expressing different antigens, in a manner to elicit immuno response in the host, directed against the different stages of the parasite's life-cycle.

In addition, the invention includes a process to produce 15 recombinant measles-malaria viruses which are avoided of defective interfering particles (DIs). The DIs are known to significantly inhibit the growth of virus in any production system and to successfully suppress immune response in human individuals.

20 Furthermore, the invention comprises a method to produce a vaccine containing such recombinant viruses.

The examples below describe the preferred mode of carrying out the invention. It should be understood that these examples are provided for illustration and should not be construed as limiting the scope of the invention in any way.

EXAMPLE 1

Construction of Recombinant MV-PfMSP-1 Plasmids

All cloning procedures were done as per the techniques described in Sambrook et al. (1989).

All the restriction enzymes were from New England BioLabs; the oligonucleotides PCR primers and DNA polylinkers were from Invitrogen.

PfMSP1 and its fragments (d-83-30-38 and d-42) either in the secreted and anchored form, have been chemically synthesized and human codon optimised. They have been cloned into the pZE21MV intermediate vector and have been slightly modified by adding SgrAI cloning site at the 5' end followed by an optimised Kozak sequence (TCATCA). These modifications have been checked by sequencing at MWG Biotech.

List of the recombinant plasmids, GPI-anchored and secreted (*) forms, from 3D7 strain, which belongs to the MAD20 prototype, and from FCB1 strain, which belongs to the K1 prototype:

p(+)MV₂EZ-d-190-SgrAI (3D7)
p(+)MV₃EZ-d-190-SgrAI (3D7)
p(+)MV₂EZ-d-83-30-38-SgrAI (3D7)
p(+)MV₃EZ-d-83-30-38-SgrAI (3D7)
p(+)MV₂EZ-d-42-SgrAI (3D7)
p(+)MV₃EZ-d-42-SgrAI (3D7)
p(+)MV₂EZ-d-190*-SgrAI (3D7)
p(+)MV₃EZ-d-190*-SgrAI (3D7)
p(+)MV₂EZ-d-83-30-38*-SgrAI (3D7)
p(+)MV₃EZ-d-83-30-38*-SgrAI (3D7)
p(+)MV₂EZ-d-42*-SgrAI (3D7)
p(+)MV₃EZ-d-42*-SgrAI (3D7)
p(+)MV₂EZ-d-190-SgrAI (FCB1)
p(+)MV₃EZ-d-190-SgrAI (FCB1)

1a) Construction of p(+)MV₂EZ-d-190-SgrAI (3D7, 24323 bp) and p(+)MV₃EZ-d-190-SgrAI (3D7, 24323 bp).
1 μg of MV plasmid DNA containing the green fluorescent protein (GFP) (p(+)MV_{2,3}EZ-GFP Berna strain, 19774 bp: FIGS. 24 and 25) was digested with one unit of both SgrAI and BssHII restriction enzymes, for two hours at their optimal

temperature, in 50 µl final volume. All the digested DNA was loaded onto a 1% agarose gel, run at 80 Volt for about 2 hours. Then, the proper band (19048 bp) was excised from the gel, purified by QIAEX gel purification and the DNA concentration was calculated by absorbance at 260 nm and adjusted to 1 µg/ml.

1 µg of d-190 gene, inserted into an intermediate plasmid (pZE21MV-d-190 SgrAI, 7564 bp.) was taken out by SgrAI-BssHII digestion (one unit of each enzyme), for two hours at their optimal temperature, in 50 µl final volume. All the digested DNA was loaded onto a 1% agarose gel, run at 80 Volt for about 2 hours. Then, the proper band (5275 bp) was excised from the gel, purified by QIAEX gel purification kit and the DNA concentration was calculated by absorbance at 260 nm and adjusted to 1 µg/ml.

Thus, the vector (MV DNA: FIG. 1) and the insert (d-190 DNA: FIG. 2), were ligated in an equimolar ratio overnight at 16°C., using one unit of T4 DNA Ligase and its own reaction buffer in 10 µl final volume.

XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard transformation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN, mini- midi and maxi kit) and restriction enzymes digestion. The right clones were sent to MWG for sequencing: the sequences, aligned with the assumed ones using a DNA Strider software, showed 100% identity.

The d-190-3D7 gene, inserted into position 2 of the MV vector (SgrAI, pos. 4060, and BssHII, pos. 9335) is represented in FIG. 4 and its Open Reading Frame (ORF) is listed in FIG. 26.

The d-190-3D7 gene, inserted into position 3 of the MV vector (SgrAI, pos. 9862, and BssHII, pos. 15137) is represented in FIG. 6.

The genome's length (starting at ACC, pos. 609, to GGT, pos. 21884) of the recombinant Measles-Malaria plasmids was a multiple of six, allowing the rescue of the recombinant MV₂₋₃-d-190-3D7 viruses.

1b) Construction of p(+)MV₂EZ-d-83-30-38-SgrAI (3D7, 23195 bp) and p(+)MV₃EZ-d-83-30-38-SgrAI (3D7, 23195 bp).

The measles vectors were prepared as detailed described in example 3a.

The pZE21MV-d-190 SgrAI was digested BstEII-AcII to cut out the d-42 fragment; a polylinker, with cohesive BstEII and AcII ends, had been ligated to obtain the intermediate plasmid pZE21MV-d-83-30-38-SgrAI (6436 bp).

The sequence of the polylinker was: 5'-GTCACCAAGCG-GCCGCAA-3'.

1 µg of pZE21MV-d-83-30-38 SgrAI was digested SgrAI-BssHII (one unit of each enzyme), for two hours at their optimal temperature, in 50 final volume. All the digested DNA was loaded onto a 1% agarose gel, run at 80 Volt for about 2 hours. Then, the proper band (4147 bp) was excised from the gel, purified by QIAEX gel purification kit and the DNA concentration was calculated by absorbance at 260 nm and adjusted to 1 µg/ml.

Thus, the vector (MV DNA: FIG. 1) and the insert (d-83-30-38 DNA: FIG. 2), were ligated in an equimolar ratio overnight at 16°C., using one unit of T4 DNA Ligase and its own reaction buffer in 10 µl final volume.

XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard transformation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN,

mini- midi and maxi kit) and restriction enzymes digestion. The right clones were sent to MWG for sequencing: the sequences were then aligned with the assumed ones using a DNA Strider software.

5 The right clones were sent to MWG for sequencing: the sequences, aligned with the assumed ones using a DNA Strider software, showed 100% identity.

The d-83-30-38-3D7 gene, inserted into position 2 of the MV vector (SgrAI, pos. 4060, and BssHII, pos. 8207) is 10 represented in FIG. 7 and its Open Reading Frame (ORF) is listed in FIG. 28.

The d-83-30-38-3D7 gene, inserted into position 3 of the MV vector (SgrAI, pos. 9862, and BssHII, pos. 14006) is 15 represented in FIG. 9.

15 The genome's length (starting at ACC, pos. 609, to GGT, pos. 20756) of the recombinant Measles-Malaria plasmids was a multiple of six, allowing the rescue of the recombinant MV₂₋₃-d-83-30-38-3D7 viruses.

1c) Construction of p(+)MV₂EZ-d-42-SgrAI (3D7, 20417 bp) and p(+)MV₃EZ-d-42-SgrAI (3D7, 20417 bp).

20 The measles vectors were prepared as detailed described in example 3a.

1 µg of d-42 gene, inserted into an intermediate plasmid (pZE21MV-d-42 SgrAI, 3658 bp) was taken out by SgrAI-BssHII digestion (one unit of each enzyme), for two hours at 25 their optimal temperature, in 50 µl final volume. All the digested DNA was loaded onto a 1% agarose gel, run at 80 Volt for about 2 hours. Then, the proper band (1369 bp) was excised from the gel, purified by QIAEX gel purification kit and the DNA concentration was calculated by absorbance at 260 nm and adjusted to 1 µg/ml.

30 Thus, the vector (MV DNA: FIG. 1) and the insert (d-42 DNA: FIG. 2), were ligated in an equimolar ratio overnight at 16°C., using one unit of T4 DNA Ligase and its own reaction buffer in 10 µl final volume.

35 XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard transformation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN, mini- midi and maxi kit) and restriction enzymes digestion. The right clones were sent to MWG for sequencing: the sequences, aligned with the assumed ones using a DNA Strider software, showed 100% identity.

40 The d-42-3D7 gene, inserted into position 2 of the MV vector (SgrAI, pos. 4060, and BssHII, pos. 5429) is represented in FIG. 10 and its Open Reading Frame (ORF) is listed in FIG. 30.

The d-42-3D7 gene, inserted into position 3 of the MV 50 vector (SgrAI, pos. 9862, and BssHII, pos. 11231) is represented in FIG. 12.

The genome's length (starting at ACC, pos. 609, to GGT, pos. 17978) of the recombinant Measles-Malaria plasmids was a multiple of six, allowing the rescue of the recombinant 55 MV₂₋₃-d-42-3D7 viruses.

The recombinant Measles-p-42 Malaria viruses and MV vaccine induced similar cytopathic effect (FIG. 36). The transgene is rather stably expressed: its expression was completely maintained in all analysed progeny clones derived from single original rescued clones after ten serial virus passages in human diploid cell MRC5 (FIG. 37-38).

The growth curves of recombinant MV-Malaria virus and MV vaccine showed the same kinetics (FIG. 39).

1d) Construction of p(+)MV₂EZ-d-190*-SgrAI (3D7, 24227 bp) and p(+)MV₃EZ-d-190*-SgrAI (3D7, 24227 bp).

60 The measles vectors were prepared as detailed described in example 3a.

15

Using the intermediate vector pZE21MVD-190-SgrAI as template, a PCR reaction has been performed to delete the GPI anchor region, which is located between AcII (pos. 5434) and ClI (pos. 5536) sites.

PCR amplifications were carried out using the proofreading Pfu DNA polymerase (Stratagene). DNA sequences of the synthetic oligonucleotides primers are given in lower case for the MV nucleotides and in upper case for non MV nucleotides; sequences of relevant restriction endonucleases recognition sites are underlined.

The following oligonucleotides primers have been used: For-ClaI, 5'-CCAATAAACGTTAAT AGAtcgattacgccgcgtctcgc-3', and Rev-AvrII, 5'-gccttgagtgaggtatacc-3'.

For-ClaI is homologous to the template at the level of the ClI and BssHII sites and contains an overhang (in upper case) with two stop codons (TAATAG), the AcII site (AACGTT), and a 6 bp long-protection site for AdI (CCAATA). In the so-called PCR-GPI and in the final construct d-190*, AcII will become close to ClI.

Rev-AvrII is homologous to the template (from pos. 5704 to 5724).

PCR product was 207 bp-long: its digestion with AcII+AvrII and ligation with the pre-digested AcII+AvrII intermediate vector pZE21MVD-190-SgrAI has produced pZE21MVD-190*-SgrAI.

In detail, the digestion of the vector with AcII-AvrII has produced two bands of 7318 bp and 246 bp (containing the GPI region to delete); the 7.3 kb-fragment was purified from agarose gel by using QIAEX II purification kit (Qiagen) and was ligated to the digested AcII-AvrII PCR (insert) to obtain pZE21MVD-190*-SgrAI.

To screen for positive clones, NcoI digestion has been done, producing a single band of 7 kb from the d-190* intermediate vector, and two bands of 1.3 and 5.7 kb from the original GPI-anchor construct.

To construct the definitive recombinant p(+)-MeV₂EZ-d190* and p(+)-MeV₃EZ-d190* (FIG. 5 and FIG. 6), according to the "rule of six", MeV vectors and intermediate plasmid were digested with SgrAI+BssHII and afterwards ligated each other.

In detail, pZE21MVD-190*-SgrAI digested SgrAI+BssHII has produced three bands, 5.2 kb+1.3 kb+900 bp. D-190* sequence was contained in the 5.2 kb fragment, that has been cut, purified and ligated with MeV₂EZ and MeV₃EZ vectors SgrAI+BssHII digested (19 Kb in length), in an equimolar ratio overnight at 16°C., using one unit of T4 DNA Ligase.

XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard transformation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN, mini- midi and maxi kit) and restriction enzymes digestion. The right clones were sent to MWG for sequencing: the sequences, aligned with the assumed ones using a DNA Strider software, showed 100% identity.

The d-190*-3D7 gene, inserted into position 2 of the MV vector (SgrAI, pos. 4060, and BssHII, pos. 9239) is represented in FIG. 5 and its Open Reading Frame (ORF) is listed in FIG. 27.

The d-190*-3D7 gene, inserted into position 3 of the MV vector (SgrAI, pos. 9862, and BssHII, pos. 15041) is represented in FIG. 6.

The genome's length (starting at ACC, pos. 609, to GGT, pos. 21788) of the recombinant Measles-Malaria plasmids was a multiple of six, allowing the rescue of the recombinant MV₂₋₃-d-190*-3D7 viruses.

16

1e) Construction of p(+)-MV₂EZ-d-83-30-38*-SgrAI (3D7, 23105 bp) and p(+)-MV₃EZ-d-83-30-38*-SgrAI (3D7, 23105 bp).

The measles vectors were prepared as detailed described in example 3a.

The intermediate vector pZE21MVD-190-SgrAI was digested BstEII-ClaI to cut out the d-42 fragment and the GPI, region, which is located between AcII (pos. 5434) and ClI (pos. 5536) sites; a polylinker, with cohesive BstEII and ClI ends, had been ligated to obtain the intermediate plasmid pZE21MV-d-83-30-38*-SgrAI (6346 bp).

The sequence of the polylinker was: 5'-GTCACCGGGGAATAATAGCGCAT-3'.

DNA sequence of the synthetic oligonucleotide polylinker is given in upper case for non MV nucleotides; sequences of relevant restriction endonucleases recognition sites are underlined.

Polylinker contains the BstEII (GTCACC) and ClI (AT) sticky ends, two stop codons (TAATAG), and a triplet (GCG) to keep the rule of six.

1 µg of pZE21MV-d-83-30-38* SgrAI was digested SgrAI-BssHII (one unit of each enzyme), for two hours at their optimal temperature, in 50 µl final volume. All the 25 digested DNA was loaded onto a 1% agarose gel, run at 80 Volt for about 2 hours. Then, the proper band (4057 bp) was excised from the gel, purified by QIAEX gel purification kit and the DNA concentration was calculated by absorbance at 260 nm and adjusted to 1 µg/ml.

Thus, the vector (MV DNA: FIG. 1) and the insert (d-83-30-38* DNA: FIG. 2), were ligated in an equimolar ratio overnight at 16°C., using one unit of T4 DNA Ligase and its own reaction buffer in 10 µl final volume.

35 XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard transformation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN, mini- midi and maxi kit) and restriction enzymes digestion. 40 The right clones were sent to MWG for sequencing: the sequences were then aligned with the assumed ones using a DNA Strider software. The right clones were sent to MWG for sequencing: the sequences, aligned with the assumed ones 45 using a DNA Strider software, showed 100% identity.

The d-83-30-38*-3D7 gene, inserted into position 2 of the MV vector (SgrAI, pos. 4060, and BssHII, pos. 8117) is represented in FIG. 8 and its Open Reading Frame (ORF) is listed in FIG. 29.

50 The d-83-30-38*-3D7 gene, inserted into position 3 of the MV vector (SgrAI, pos. 9862, and BssHII, pos. 13919) is represented in FIG. 9.

The genome's length (starting at ACC, pos. 609, to GGT, pos. 20666) of the recombinant Measles-Malaria plasmids 55 was a multiple of six, allowing the rescue of the recombinant MV₂₋₃-d-83-30-38*-3D7 viruses.

1f) Construction of p(+)-MV₂EZ-d-42*-SgrAI (3D7, 20345 bp) and p(+)-MV₃EZ-d-42*-SgrAI (3D7, 20345 bp).

The measles vectors were prepared as detailed described in example 3a.

Using the intermediate vector pZE21MVD-42-SgrAI (3658 bp) as template, a PCR reaction has been performed to delete the GPI anchor region, which is located between AcII (pos. 1528) and ClI (pos. 1630) sites.

65 PCR amplifications were carried out using the proofreading Pfu DNA polymerase (Stratagene). DNA sequences of the synthetic oligonucleotides primers are given in lower case for

the MV nucleotides and in upper case for non MV nucleotides; sequences of relevant restriction endonucleases recognition sites are underlined.

The following oligonucleotides primers have been used: For-ClaI, 5'-CCAATAAACGTTTAAT AGatcgattac ggcgctttagc-3', and Rev-AvrII, 5'-gcctttaggtgagctgatacc-3'.

For-ClaI is homologous to the template at the level of the ClaI (pos. 1630) and BssHII (pos. 1639) sites and contains an overhang (in upper case) with two stop codons (TAATAG), the AclI site (AACGTT), and a 6 bp long-protection site for AclI (CCAATA). In the so-called PCR-GPI and in the final construct d-42*, AclI will become close to ClaI. Rev-AvrII is homologous to the template (from pos. 1798 to 1818).

PCR product was 207 bp-long: its digestion with AclI+AvrII and ligation with the pre-digested AclI+AvrII intermediate vector pZE21MVD-42-SgrAI has produced pZE21MVD-42*-SgrAI.

In detail, the digestion of the vector with AclI+AvrII has produced two bands of 3412 bp and 246 bp (containing the GPI region to delete): the 3.4 kb-fragment was purified from agarose gel by using QIAEX II purification kit (Qiagen) and was ligated to the digested AclI-AvrII PCR (insert) to obtain pZE21MVD-42*-SgrAI.

To screen for positive clones, NcoI digestion has been done, producing a single band of 3.4 kb from the d-42* intermediate vector, and two bands of 1.3 and 2.3 kb from the original GPI-anchor construct.

To construct the definitive recombinant p(+)-MeV₂EZ-d42* and p(+)-MeV₃EZ-d42*, according to the "rule of six", MeV vectors and intermediate plasmid were digested with SgrAI+BssHII and afterwards ligated each other.

In detail, pZE21MVD-42*-SgrAI digested SgrAI+BssHII+SpeI has produced four bands, 1.3 kb+936 bp+800 bp+400 bp. D-42* sequence was contained in the 1.3 kb fragment, that has been cut, purified and ligated with MeV₂EZ and MeV₃EZ vectors SgrAI+BssHII digested (19 Kb in length), in an equimolar ratio overnight at 16°C., using one unit of T4 DNA Ligase.

XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard transformation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN, mini- midi and maxi kit) and restriction enzymes digestion. The right clones were sent to MWG for sequencing: the sequences, aligned with the assumed ones using a DNA Strider software, showed 100% identity.

The d-42*-3D7 gene, inserted into position 2 of the MV vector (SgrAI, pos. 4060, and BssHII, pos. 5357) is represented in FIG. 11 and its Open Reading Frame (ORF) is listed in FIG. 31.

The d-42*-3D7 gene, inserted into position 3 of the MV vector (SgrAI, pos. 9862, and BssHII, pos. 11159) is represented in FIG. 12.

The genome's length (starting at ACC, pos. 609, to GGT, pos. 17906) of the recombinant Measles-Malaria plasmids was a multiple of six, allowing the rescue of the recombinant MV₂₋₃-d-42*-3D7 viruses.

1g) Construction of p(+)-MV₂EZ-d-190-SgrAI (FCB1, 24083 bp) and p(+)-MV₃EZ-d-190-SgrAI (FCB1, 24083 bp).

First of all, the cloning of the synthetic gene for MSP-1 of the FCB1 strain into the intermediate plasmid pZE21MV-SgrAI has been performed, keeping the signal peptide and the GPI-anchor region from MSP-1 of 3D7 strain. D-190 gene (FCB1) was obtained stepwise from an intermediate vector, called pZE23f-GX-190H, as follow:

i). 1 µg of the plasmid pZE21MV-d-190-SgrAI (3D7) was digested with HindIII+AclI restriction enzymes, for two hours at their optimal temperature, in 50 µl final volume. All the digested DNA was loaded onto a 1% agarose gel, run at 80 Volt for about 2 hours. Then, the proper band (2558 bp), corresponding to the vector, was excised from the gel, purified by QIAEX gel purification and the DNA concentration was calculated by absorbance at 260 nm.

ii). a PCR reaction was performed, using the pZE23f-GX-190H as template, in order to amplify and recover the d-42 portion of the MSP-1/FCB1. PCR amplification was carried out using the proofreading Pfu DNA polymerase (Stratagene). DNA sequences of the synthetic oligonucleotides primers are given in lower case for the MV nucleotides and in upper case for non MV nucleotides; sequences of relevant restriction endonucleases recognition sites are underlined.

The following oligonucleotides primers have been used, designed on the pZE23f-GX-190H sequence: For-1 FCB1, 5'-CCCAAGCTTccaggtagtcaccggAgagctgtcactcc-3', and Rev-1 FCB1, 5'-GCCTGCaacgttGCTagagctggagcaGaaGatcccgatcg-3'.

For-1 FCB1 is homologous to the template from pos. 4509 to pos. 4538, comprising the BstEII site (ggtcacc). The A (in upper case) was a t in the template, and it has been modified to eliminate a SgrAI site. It contains an overhang (in upper case) with the HindIII site (AAGCTT), after its 3 bp long-protection site (CCC).

Rev-1 FCB1 contains an AclI site (aacgtt), preceded by a 6-bp protection site (GCCTGC). It was introduced a triplet GCT, coding for a serine, to keep the rule of six; two a have been modified in G to avoid a poly(A) site.

The obtained PCR-HindIII-AclI (1.1 kb) has been digested HindIII+AclI and ligated, overnight at 16°C. in an equimolar ratio, to the pre-digested pZE21MV-d-190-SgrAI with HindIII+AclI (step i), obtaining the pZE21MV-d-42-SgrAI-FCB1 (3657 bp). XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard transformation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN, mini- midi and maxi kit) and by restriction enzymes digestion with HindIII+AclI (expected fragments 2558 bp+1099 bp).

iii). the pZE21MV-d-42-SgrAI-FCB1, obtained as described in step ii, has been digested HindIII+BstEII (HindIII, pos. 428, and BstEII, pos. 440), and the proper band (3645 bp), corresponding to the opened vector, was loaded on a 1% agarose gel, excised from the gel, purified by QIAEX gel purification and the DNA concentration was calculated by absorbance at 260 nm.

iv). The pZE23f-GX-190H was digested HindIII+BstEII and the proper band of 3679 bp (insert), corresponding to the d-83-30-38/FCB1 fragment, was purified from the gel, as previously described.

v). The HindIII+BstEII digested fragment of 3657 bp (vector), obtained from pZE21MV-d-42-SgrAI-FCB1, has been ligated to the HindIII+BstEII fragment of 3679 bp (insert), containing the d-83-30-38/FCB1 and obtained by digestion from pZE23f-GX-190H. Ligation was done in an equimolar ratio overnight at 16°C., using one unit of T4 DNA Ligase, obtaining the pZE21MV-d-190-SgrAI-FCB1 (7324 bp). Afterwards, XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard transformation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN, mini- midi and maxi kit) and restriction enzymes digestion.

19

To construct the p(+)MV₂EZ-d-190-SgrAI-FCB1 and p(+)MV₃EZ-d-190-SgrAI-FCB1, the measles vectors were prepared as detailed described in example 3a.

1 µg of d-190/FCB1 gene, inserted into an intermediate plasmid (pZE21MV-d-190 SgrAI-FCB1, 7324 bp), was taken out by SgrAI+BssHII digestion (one unit of each enzyme), for two hours at their optimal temperature, in 50 µl final volume. All the digested DNA was loaded onto a 1% agarose gel, run at 80 Volt for about 2 hours. Then, the proper band (5035 bp) was excised from the gel, purified by QIAEX gel purification kit and the DNA concentration was calculated by absorbance at 260 nm and adjusted to 1 µg/ml.

Thus, the vector (MV DNA: FIG. 1) and the insert (d-190/FCB1 DNA: FIG. 3), were ligated in an equimolar ratio overnight at 16°C, using one unit of T4 DNA Ligase and its own reaction buffer in 10 µl final volume.

XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard transformation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN, mini- midi and maxi kit) and restriction enzymes digestion. The right clones were sent to MWG for sequencing: the sequences, aligned with the assumed ones using a DNA Strider software, showed 100% identity.

The d-190-FCB1 gene, inserted into position 2 of the MV vector (SgrAI, pos. 4060, and BssHII, pos. 9095) is represented in FIG. 13 and its Open Reading Frame (ORF) is listed in FIG. 32.

The d-190-FCB1 gene, inserted into position 3 of the MV vector (SgrAI, pos. 9862, and BssHII, pos. 14897) is represented in FIG. 14.

The genome's length (starting at ACC, pos. 609, to GGT, pos. 21884) of the recombinant Measles-Malaria plasmids was a multiple of six, allowing the rescue of the recombinant MV_{2,3}-d-190-FCB1 viruses.

The transgene is rather stably expressed: its expression was completely maintained in all analysed progeny clones derived from single original rescued clones after ten serial virus passages in human diploid cell MRC5 (FIG. 40).

The growth curves of recombinant MV-Malaria virus and MV vaccine showed the same kinetics (FIG. 41).

EXAMPLE 2

Designing of DiCo1 Nucleic Acid Sequence

Starting from the aminoacidic DiCo1 sequence (ecto, trans and cytoplasmic domains: aa 97-622) and using the DNA Strider software, a correspondent nucleic acid sequence has been designed comparing the DiCo1 DNA degenerate sequence to a selected PfAMA1 gene (accession number AAG141.1), which represents the most similar sequence to the DiCo1 after BLAST alignment.

At the 5' end suitable unique restriction sites has been added (MluI and SgrAI) as cloning sites, followed by an optimal KOZAC sequence and a human optimised Signal Peptide (SP). At the 3' end, two stop codons and a BssHII cloning site have been added. Following this scheme, we designed two nucleotides sequences (respecting the "rule of six" for the further expression into the measles vector), encoding the anchored and the secreted forms of the DiCo1 protein: the first gene comprises the ectoplasmasic, the transmembrane and cytoplasmic domains (FIG. 18), while the second one corresponds to the ectodomain alone (FIG. 19). The two sequences has been human codon optimised by

20

GENEART, to reduce AT % content, to avoid poly(A) sequence and RNA instability motif.

DiCo1 complete ORF and DiCo1 ectodomain ORF are listed respectively in FIGS. 34 and 35.

EXAMPLE 3

Construction of Recombinant MV-PfAMA-1 Plasmids

All cloning procedures were done as per techniques described in Sambrook et al. (1989).

PfAMA1, and in particular Diversity Covering sequences 1 (DiCo1) either in the secreted and anchored form, have been chemically synthesized and human codon optimised.

The codon optimised DiCo1 secreted and anchored forms were digested SgrAI+BssHII and ligated, overnight at 16°C, in an equimolar ratio, to the pre-digested MeV₂EZ and MeV₃EZ vectors (19 Kb in length), using one unit of T4 DNA Ligase, obtaining the following recombinant MV-PfAMA-1 plasmids: p(+)MV₂EZ-DiCo1-complete (FIG. 20), p(+)MV₃EZ-DiCo1-complete (FIG. 21), p(+)MV₂EZ-DiCo1-ecto (FIG. 22), and p(+)MV₃EZ-DiCo1-ecto (FIG. 23).

EXAMPLE 4

Construction of Recombinant MV-PJCS Plasmids

Construction of p(+)MV₂EZ-CS-SgrAI (20219 bp) and p(+)MV₃EZ-CS-SgrAI (20219 bp)

All cloning procedures were basically as described in Sambrook et al. (1989).

PJCS1, cloned into an intermediate vector pAdApt35BsU.CS.Pfalc.aa-sub.gcc, has been amplified by PCR, and directly cloned into the definitive MV vectors, obtaining two recombinant MV-PJCS plasmids: p(+)MV₂EZ-CS and p(+)MV₃EZ-CS.

In detail, a PCR reaction was performed, using the pAdApt35BsU.CS.Pfalc.aa-sub.gcc as template, in order to amplify and recover the CS gene (FIG. 15). PCR amplification was carried out using the proofreading Pfu DNA polymerase (Stratagene). DNA sequences of the synthetic oligonucleotides primers are given in lower case for the MV nucleotides and in upper case for non MV nucleotides; sequences of relevant restriction endonucleases recognition sites are underlined.

The following oligonucleotides primers have been used, designed on the pAdApt35BsU.CS.Pfalc.aa-sub.gcc sequence: For-SgrAI, 5'-ACTTCTCACCGGTGTgg aagcttgcac catgtat-3', and Rev-BssHII-CS 5'-TA GCGCGCtctagaggatcttcgc-3'.

For-SgrAI is homologous to the template from pos. 1356 to pos. 1375, comprising the HindIII site (aagctt). It contains an overhang (in upper case) with SgrAI restriction site (CAC-55 CGGTG), after 6-bp long-protection site (ACTTCT).

Rev-BssHII-CS contains an overhang (in upper case) with BssHII restriction site (GCGCGC), which will be close to XbaI (tctaga) in the PCR-CS (1187 bp).

The obtained PCR-CS has been digested SgrAI+BssHII and ligated, overnight at 16°C, in an equimolar ratio, to the pre-digested MeV₂EZ and MeV₃EZ vectors SgrAI+BssHII (19 Kb in length), using one unit of T4 DNA Ligase, obtaining, respectively, p(+)MV₂EZ-CS-SgrAI (20219 bp, FIG. 16) and p(+)MV₃EZ-CS-SgrAI (20219 bp, FIG. 17). The CS ORF is listed in FIG. 33.

XL10 Gold chemical competent cell were then transformed with all ligation volume, following a standard trans-

21

formation protocol (Sambrook et al. 1989), plated and selected on LB-Agar plates for ampicillin resistance. Colonies were screened by DNA plasmid preparation (QIAGEN, mini- midi and maxi kit) and restriction enzymes digestion. The right clones were sent to MWG for sequencing: the sequences, aligned with the assumed ones using a DNA Strider software, showed 100% identity.

EXAMPLE 5

Cells and Viruses

Cells were maintained as monolayers in Dulbecco's Modified Eagles Medium (DMEM), supplemented with 5% Foetal Calf Serum (FCS) for Vero cells (African green monkey kidney) and with 10% FCS and 1% penicillin/streptomycin (P/S) for 293T cells (human embryonic kidney); DMEM supplemented with Glutamax (F12) and 10% FCS for MRC-5 (human foetal fibroblast); DMEM supplemented with 10% FCS and 1.2 mg/ml of G 418 for 293-3-46.

To grow MV virus stocks reaching titers of about 10⁷ pfu/ml, recombinant viruses and the vaccine strain Edmoston Zagreb were propagated in MRC-5 cells; plaque purification was carried out by transferring a syncytium to 35 mm MRC-5 cell culture which was expanded first to a 10 cm dish, and afterwards to a 175 cm² flask. Virus stocks were made from 175 cm² cultures when syncytia formation was about 90% pronounced. Medium corresponding to the so-called "free-cell virus fraction" was collected, freeze and thawed three times and spun down to avoid cell debris. The medium was then stored at -80° C. Cells, which correspond to the so-called "cell-associated virus fraction", were scraped into 3 ml of OPTIMEM (Gibco BRL) followed by three rounds freezing and thawing, spun down and the cleared supernatant stored at -80° C.

EXAMPLE 6

Transfection of Plasmids and Rescue of MV Viruses

293T cells were seeded into a 35 mm well to reach ~50-70% confluence when being transfected. 4 h before transfection, the medium was replaced with 3 ml DMEM containing 10% FCS. All recombinant plasmids were prepared according to the QIAGEN plasmid preparation kit. The kit for the Ca²⁺ phosphate coprecipitation of DNA was from Invitrogen.

Cells were co-transfected with the plasmids in the follows final concentration: pCA-L 0.5 µg, pCA-N 0.5 µg, pCA-P 0.1 µg, pCA T7 1 µg and the recombinant Measles-Malaria plasmid 4 µg. All five plasmids, diluted in H₂O, were added in a Eppendorf tube containing 2M CaCl₂, the mix was added to another Eppendorf tube containing HEPES buffer under shaking conditions, and was incubated 30 min at room temperature (RT). Thus, the co-precipitates were added dropwise to the culture and the transfection was carried out at 37° C. and 5% CO₂ for about 18 h. Then, the transfection medium was replaced with 3 ml of DMEM containing 10% FCS.

Another way to obtain recombinant measles-malaria vaccine viruses is described hereafter, using the 293-3-46 helper cell (human embryonic kidney cells), stably expressing the measles N and P proteins as well as the T7 RNA polymerase. The viral RNA polymerase (large protein, L) was expressed by co-transfecting the cells with 15 ng of the plasmid pM-CLa. To improve transfection efficiency 300 ng of pSC6-Neo were added. Calcium-phosphate method was used for transfection.

22

First syncytia appeared 3-4 days after transfection when the cells were still subconfluent. To allow syncytia formation to progress more easily, almost confluent cell monolayer of each 35 mm well were then transferred to a 10 cm dish. Each syncytium was taken up in 300 µl of transfection medium and put in a sterile Eppendorf tube containing 700 µl of OPTIMEM, freeze and thaw for three rounds, and stored at -80° C.

EXAMPLE 7

Virus Titration by Plaque Assay

Serial 10-times dilutions of virus preparations were carried out using OPTIMEM to a final volume of 0.5 ml. Each dilution was added on 35 mm Vero cell cultures. After 1 h of virus adsorption, the inoculum was removed and the infected cells were overlaid with 2 ml of DMEM containing 5% FCS and 1% low melting point agarose (LMP agarose). After 5 days of incubation at 37° C. and 5% CO₂, cultures were fixed with 1 ml of 10% TCA for 1 h, then UV cross-linked for 30 min. After removal of the agarose overlay, cell monolayers were stained with crystal violet dissolved in 4% ethanol, washed with water and the plaques were counted under the inverted microscope.

EXAMPLE 8

MRC-5 Virus Serial Passages of Recombinant Viruses

Rescued viruses were serially passaged 10-times on MRC5 cells, seeded into 10 cm diameter plates, that were infected with the standard and the recombinant MV viruses at MOI of 0.01 PFU/cells. After monolayer was full infected, 1% supernatant of each culture was used to infect the subsequent MRC5 cells monolayer. To test transgene expression and stability, viruses from passage 1, 5, and 10 were used for further characterisation of expression by Western blot and immunofluorescence.

EXAMPLE 9

Western Blot, Immunofluorescence

To analyse the expression either MV and Malaria, Western blot and immunofluorescence were carried out.

For Western blot, Vero cells seeded on 35 mm dish (1-5×10⁵) were monitored the next day for 90% confluence and infected with cleared virus suspension from cell-associated virus fraction, using 0.1 MOI (Multiplicity Of Infection), including MVEZ as control. When about 80% syncytia formation was observed, cells were first washed with PBS and then scraped in 1 ml PBS and collected in an Eppendorf tube, and centrifuge at 2000 RPM/4 min. Cells were then lysated 5 min/RT with 70 µl of lysis buffer (1% NP-40, 50 mM Tris pH 8, 150 mM NaCl) supplemented with protease inhibitor cocktail (Complete Mini, Roche, 1 836 153). Supernatants were cleared by centrifuge at 13000 RPM/5 min, and transferred into a new tube: 30 µl of 4× loading buffer (Invitrogen) were added; samples were mixed and boiled at 95° C./2 min, spun down and stored at -20° C.

An SDS-PAGE migration was performed, running a NuPAGE 12% Bis-acrylamide gel in reducing conditions, using 1× Running Buffer, for 50 min at 200V (start 100-125 mA, end 60-80 mA).

Then, semi-dry method was used to transfer separated cell proteins to Nitrocellulose Membrane, at 14V/1 h 30.

23

As first antibodies, rabbit polyclonal against MSP1-p-83, diluted in PBST at least 1:30000, and against MSP1-p-42,* diluted at least 1:50000, were used. The second antibody was a swine anti-rabbit antibody coupled to horse-radish peroxidase allowing the visualization of the bands by the enhanced chemiluminescence kit (ECLTM, Amersham LifeScience).

For immunofluorescence, Vero cells were seeded on a 24 mm×24 mm glass cover slips in 35 mm wells, cultured overnight and infected with rescued recombinant virus. 3 days after infection cells on coverslips were fixed with 3.7% paraformaldehyde in PBS, and permeabilized with 0.1% TX-100, washed with blocking solution (PBS containing 1% BSA) for 1 h, and stained with the specific antibodies. Mouse hybridoma supernatant mAb 5.2, which recognises a EGF-like domain in the p-19 portion of p-42, was used in a dilution 1:100 followed by FITC conjugated goat anti-mouse serum, diluted 1:250.

EXAMPLE 10

Growth Kinetics Curve

MRC5 cells seeded on 35 mm dish ($1\text{-}5 \times 10^5$) were monitored for 90% confluence and infected with cleared virus suspension from cell-associated virus fraction, using 0.1 MOI, including MVEZ as control. Samples, corresponding to the so-called “free-cell virus fraction” and to the so-called “cell-associated virus fraction”, were collected daily for one week and titrated.

EXAMPLE 11

Mice Immunisation

The immunogenic power of the rescued recombinant MV-Malaria viruses described was proven by immunisation tests performed on transgenic mice IFNAR/CD46, susceptible to MV infections. The animals were kept under optimal hygienic conditions and were immunized at 6-8 weeks of age. Below is provided an example of mice immunization with two recombinant Measles-Malaria virus: the MeV2EZ-d-p42-SgrAI (the GPI anchored form) and the MeV2EZ-d-p42* (the secreted form). Immunisation was performed intramuscularly using 10^5 PFU of each recombinant MV-Malaria in three injections at 0, 4 and 8 weeks. Mice immunized with recombinant-empty Measles (rMVEZ13-Empty cloned) served as negative control. UV inactivated rMV was used as a control to determine the effect of virus replication on activation of immune responses. The immune response of the MV vectored antigen was tested compared to the purified d-42 protein (0.5 mg/ml): mice were immunized sub cutaneously with 20 µg of protein in Incomplete Freund's Adjuvant.

Blood was taken regularly and tested for Measles IgG Titers (FIG. 42).

Blood was taken regularly and tested for Malaria IgG Titers (FIG. 43).

The presence of MV-specific antibodies in the sera from the immunised IFNAR/CD46 mice (6 per test group and 3 for control group) was determined by ELISA using 96-microwell plates, coated with Measles virus EIA bulk (ATCC VR-24), for IgG antibody detection. Protein was diluted 0.6 µg/ml with 0.05 M carbonate buffer (pH 9.4), and 100 µl per well was added to 96-well-microtiter plates. The plates were incubated overnight at 4° C., washed with PBS/0.05% Tween 20 (PT) (ph 7.4), incubated with PT (0.1 ml/well)-10% BSA for 60 min at 37° C., and washed again with PT. Serial 2-folds dilutions of the tested sera were added (100 µl/well), and the

24

plates were incubated for 60 min at 37° C. The plates were washed with PT and were incubated with 100 µl of goat anti-mouse IgG HRP diluted 1:2000 in PT for 30 min at 37° C. The plates were washed with PT and incubated with 100 µl OPD (o-Phenylenediamin, Fluka 78411). The reaction was stopped after 3-4 min. Plates were read on a MicroElisa Reader at a wave length of 490 nm. Readings higher than three-folds negative controls were scored as positive reaction.

The presence of MV-Malaria-specific antibodies in the sera of immunised CD46 mice (at least 10 per test group) was determined by ELISA assay. Briefly, 96-microwell plates were coated 50 ng/well MSP-1-d42 3D7 strains, diluted with carbonate buffer pH 9.4. The plates were incubated overnight at 4° C., washed with PBS/0.05% Tween 20 (PT). Subsequently, unspecific interaction were blocked with 10% defatted milk dissolved in PT for 1hour at 37° C. and wells were washed again with PT. The plates were consecutively incubated with various dilutions of mouse sera (starting at 1:200, followed by serial two-fold dilutions), peroxidase-conjugate goat anti-mouse IgG and with OPD substrate. Optical density values were measured at 490 nm. Values above the cut-off background level (mean value of sera from MV immunised mice multiplied by a factor of 2.1) were considered positive. Titres were depicted as reciprocal end-dilutions.

The humoral immune responses against Measles are shown in FIG. 42. The humoral immune responses against Malaria p42 are shown in FIG. 43.

EXAMPLE 12

Purification of Recombinant Measles Virus Expressing Malaria Antigens from Defecting Interfering Particles (DIs) by Plaque Purification

It is known from literature that after a certain number of passages with Paramyxoviruses, and in particular with measles virus, an accumulation of defective interfering particles (DIs) will occur (23, 24). It has been described that these DIs develop various defects: negative impact on vaccine safety, negative influence on virus yields in production, genome instability and suppression of immune reaction after vaccination. In order to avoid such DIs with our new recombinant viruses, we have applied the method of plaque purification as described in example 6 with the exception that we use MRC5 cell instead of 293T cells. After the formation of clear, well defined syncytia we aspirated under the microscope with a micropipette such material for further passaging in a fresh MRC5 tissue culture.

EXAMPLE 13

Purification of Recombinant Measles Virus Expressing Malaria Antigens from Defecting Interfering Particles (DIs) by End Point Dilution

The end point dilution technique was applied in microplates: in all wells a fresh monolayer of MRC5 cells had just developed. The virus suspension containing recombinant measles-malaria viruses was prepared in two fold dilutions. From the well of the latest monolayer where a syncytia was detected the supernatant was aspirated with a pipette. The supernatant was mixed with a suspension containing MRC5 cells. This mixture was incubated at 4° C. for 1 hour. Finally, it was transferred in a small Costar flask and incubated at 35° C./5% CO₂ and harvested for purify recombinant measles-malaria virus after ten days.

Production of a Combined Measles-Malaria Vaccine

The working seed of the described recombinant measles-malaria virus has been incubated on MRC5 cell monolayer in 1750 cm² roller bottles at 35° C. for ten days. The cells have been monitored every day for status of health and confluence. On day ten at highest level of syncytia formation, the supernatant was pumped in a steel cylinder for storage in liquid nitrogen. The same procedure was repeated two days later. After performing of all the tests (virus titer, genome stability, virus safety, cell safety, chemical analysis, sterility and others), the harvests have been thawed up and mixed with stabilizer containing gelatine, sorbitol, aminoacids and other sugars to final dilution of 10⁵. With a automated filling machine small lyo bottles (F3) have been inoculated with 0.5 ml each. A specially calculated lyophilisation program was used to guarantee maximal survival of the product during the freeze-drying process.

BIBLIOGRAPHY

1. Fields Virology, fifth edition (2007), eds.-in-chief Knipe, D. M. & Howley, P. M. Lippincott Williams & Wilkins, Philadelphia Pa. 19106, USA.
2. Enders, J. F., and Peebles, T. C. (1954). Propagation in tissue cultures of cytopathogenic agents from patients with measles. Proc. Soc. Exp. Biol. Med., 86: 277-286.
3. Griffin, D. (2007) Measles virus. In: Fields Virology, fifth edition, eds.-in-chief Knipe, D. M. & Howley, P. M. Lippincott Williams & Wilkins, Philadelphia Pa. 19106, USA.
4. Parks, C. L., Lerch, R. A., Walpita, P., Wang, H. P., Sidhu, M. S., and Udem, S. A. (2001). Analysis of the noncoding regions of measles virus strains in the Edmonton vaccine lineage. J. Virol., 75: 921-933.
5. Parks, C. L., Lerch, R. A., Walpita, P., Wang, H. P., Sidhu, M. S., and Udem, S. A. (2001). Comparison of predicted amino acid sequences of measles virus strains in the Edmonston vaccine lineage. J. Virol., 75: 910-920.
6. Hilleman, M. R. (2002). Current overview of the pathogenesis and prophylaxis of measles with focus on practical implications. Vaccine, 20: 651-665.
7. Ovsyannikova I G., Reid, K. C., Jacobson, R. M., Oberg, A. L., Klee, G. G., Poland, G. A. (2003). Cytokine production patterns and antibody response to measles vaccine. Vaccine, 21(25-26): 3946-53.
8. Radecke, F., P. Spielhofer, H. Schneider, K. Kaelin, M. Huber, K. Dötsch, G. Christiansen, and M. Billeter. (1995). Rescue of measles viruses from cloned DNA. EMBO Journal., 14: 5773-5784.
9. Martin, A., Staeheli, P. and Schneider, U. (2006). RNA polymerase II-controlled expression of antigenomic RNA enhances the rescue efficacies of two different members of the Mononegavirales independently of the site of viral genome replication. J. Virol., 80:5708-5715.
10. Radecke, F., and M. Billeter. (1997). Reverse genetics meets the nonsegmented negative-strand RNA viruses. Rev. Med. Virol., 7: 49-63.
11. Singh M. R., Cattaneo, R., Billeter, M. A. (1999). A recombinant measles virus expressing hepatitis B virus surface antigen induces humoral immune responses in genetically modified mice. J. Virol., 73: 4823-4828.
12. Wang, Z. L., Hangartner, L., Cornu, T. I., Martin, L. R., Zuniga, A., and Billeter, M. (2001). Recombinant measles viruses expressing heterologous antigens of mumps and simian immunodeficiency viruses. Vaccine, 19: 2329-2336.
13. Dilraj, A., Cutts, F. T., de Castro, J. F., Wheeler, J. G., Brown, D., Roth, C., Coovadia, H. M., Bennett, J. V. (2000). Response to different measles vaccine strains given by aerosol and subcutaneous routes to schoolchildren: a randomised trial. Lancet, 355(9206): 798-803.
14. Holder A. A. and Freeman, R. R. (1984). The three major antigens on the surface of *Plasmodium falciparum* merozoites are derived from a single high molecular weight precursor. J. Exp. Med., 160(2): 624-9.
15. Blackman M. J., Whittle H., and Holder A. A. (1991). Processing of the *Plasmodium falciparum* major merozoite surface protein-1: identification of a 33-kilodalton secondary processing product which is shed prior to erythrocyte invasion. Mol. Biochem. Parasitol., 49(1): 35-44.
16. Remarque, E. J., Faber, B. W., Kocken, C. H. M., and Thomas, A. W. (2007). Apical membrane antigen 1: a malaria vaccine candidate in review. Trends Parasitol, 24: 74-83.
17. Polley, S. D., Mwangi, T., Kocken, C. H., Thomas, A. W., Dutta, S., Lanar, D. E., Remarque, E., Ross, A., Williams, T. N., Mwambingu, G., Lowe, B., Conway, D. J., and Marsh, K. (2004). Human antibodies to recombinant protein constructs of *Plasmodium falciparum* apical membrane antigen 1 (AMA1) and their association with protection from malaria. Vaccine, 23: 718-728.
18. Cortés, A., Mellombo, M., Masciantonio, R., Murphy, V. J., Reeder, J. C., and Anders, R. F. (2005). Allele specificity of naturally acquired antibody responses against *Plasmodium falciparum* apical membrane antigen 1. Infect. Immun., 73: 422-430.
19. Remarque, E. J., Faber, B. W., Kocken, C. H. M., and Thomas, A. W. (2008). A diversity-covering approach to immunisation with *Plasmodium falciparum* AMA1 induces broader allelic recognition and growth inhibition responses in rabbits. Infect. Immun.
20. Garcia, J. E., Fuentes, A., and Patarroyo, M. E. (2006). Developmental biology of sporozoite-host interactions in *Plasmodium falciparum* malaria: implications for vaccine design. Clin. Microbiol. Rev., 19(4): 686-707.
21. Ballou, W. R., and Cahill, C. P. (2007). Two Decades of Commitment to Malaria Vaccine Development: GlaxoSmithKline Biologicals. Am. J. Trop. Med. Hyg., 77(6_Suppl): 289-295.
22. Girard, M. P., Reed, Z. H., Friede, M., and Kieny, M. P. (2007). A review of human vaccine research and development: Malaria. Vaccine, 25: 1567-1580.
23. Roux, L., Simon, A. E., Holland, J. J. (1991). Effects of Defective Interfering Viruses on virus replication and pathogenesis in vitro and in vivo. Adv. Virus Res., 40: 181-221.
24. Calain, P., and Roux, L. (1988). Generation of measles virus defective interfering particles and their presence in a preparation of attenuated live-virus vaccine. J. Virol., 62 (8):2859-2866.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 12

<210> SEQ ID NO 1
<211> LENGTH: 19793
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Complete nucleotide sequence of p(+)MV2EZ-GFP

<400> SEQUENCE: 1

cacctaatt gtaagcgtta atatttgtt	aaaattcgcg ttaaatttt	gttaaatcag	60			
ctcattttt aaccaatagg ccgaatcg	caaataccct tataaatcaa	aagaatagac	120			
cgagataggg ttgagtgtt	ttccagttt	gaacaagagt ccactattaa	agaacgtgga	180		
ctccaaacgtc aaaggcgaa aaaccgtcta	tcagggcgat ggcccactac	gtgaaccatc	240			
accctaatac agtttttgg	ggtcgaggtg	ccgtaaagca ctaaatcgga	accctaaagg	300		
gagcccccgta	tttagagctt	gacggggaaa gccggccatt	taggccatag ggcgctggca	360		
agtgttagcggt	tcacgctgct	cgtaaccacc acaccgcgc	cgcttaatgc gccgctacag	420		
ggcgctccc attcgccatt	caggctgctc	aactgttggg aagggcgatc	ggtgccggcc	480		
tcttcgtat taaggccagct	ggcgaagggg	ggatgtgctg	caaggcgatt aagttggta	540		
acgccagggt tttccagtc	acgacgttgt	aaaacgacgg ccagtgaatt	gtaatacgc	600		
tcactataac caaacaaagt	tggtaagga	tagtcaatc aatgatcatc	ttcttagtgca	660		
cttaggattc aagatcctat	tatcaggggac	aagagcgagga	ttagggatat ctgagatggc	720		
cacacttta aggagcttag	cattgttcaa	aagaaacaag	gacaaaccac ccattacatc	780		
aggatccgggt	ggagccatca	gaggaatcaa	acacattatt atagtaccaa	tccctggaga	840	
ttcctcaatt accactcgat	ccagacttct	ggaccggttt	gtcaggttaa ttggaaaccc	900		
ggatgtgagc	ggggccaaac	taacaggggc	actaataggt atattatcct tatttggta	960		
gtctccaggt	caattgattc	agaggatcac	cgatgaccct gacgttagca	taaggctgtt	1020	
agagggttgc	cagagtgacc	agtccaaatc	tggccttacc ttgcataa	gaggtagcaa	1080	
catggaggat	gagggggacc	aatactttc	acatgatgat	ccaatttagta	gtgatcaatc	1140
cagggttgcga	tggttcgaga	acaaggaaat	ctcagatatt	gaagtgcag	accctgaggg	1200
attcaacatg	attctggta	ccatcctagc	ccaaatttgg	gtcttgctcg	caaaggcggt	1260
tacggcccca	gacacggcag	ctgattcgga	gctaagaagg	tggataaagt	acacccaaca	1320
aagaagggtt	gttggtaat	ttagattgga	gagaaaatgg	ttggatgtgg	tgaggaacag	1380
gattgecgag	gacctctcct	tacgcccatt	catggtcgt	ctaatectgg	atatcaagag	1440
aacacccgga	aacaaaccca	ggattgtcga	aatgatatgt	gacattgata	catatatcgt	1500
agaggcagga	ttagccagtt	ttatcctgac	tattaagttt	gggatagaaa	ctatgtatcc	1560
tgctcttggaa	ctgcatgaat	ttgctggtga	gttatccaca	cttgagtct	tgtgaacct	1620
ttaccagcaa	atggggaaa	ctgcaccata	catggtaatc	ctggagaact	caattcagaa	1680
caagttcaat	gcaggatcat	accctctgt	ctggagctat	gccccatggag	taggatgtgg	1740
acttgaaaac	tccatggggg	gtttgaactt	tggccgatct	tactttgatc	cagcatattt	1800
tagattaggg	caagagatgg	taaggaggc	agctggaaag	gtcagttcca	cattggcatc	1860
tgaactcggt	atcaactgccc	aggatgcaag	gcttgttca	gagattgcaa	tgcataactac	1920
tgaggacaag	atcagtagag	cgggtggacc	cagacaagcc	caagttatcat	ttctacacgg	1980
tgtatcaaagt	gagaatgagc	taccgagatt	ggggggcaag	gaagatagga	gggtcaaaaca	2040

-continued

gagtcgagga gaagccaggg agagctacag agaaaccggg cccagcagag caagtatgc 2100
 gagagctgcc catcttccaa ccggcacacc cctagacatt gacactgcat cgaggatccag 2160
 ccaagatccg caggacagtc gaaggtcage tgacgcctg cttaggctgc aagccatggc 2220
 aggaatctcg gaagaacaag gctcagacac ggacaccccct atagtgtaca atgacagaaa 2280
 tcttctagac taggtgcgag aggccgaggg ccagaacaac atccgcctac cctccatcat 2340
 ttttataaaa aacttaggaa ccaggatccac acagccgcca gcccatcaac catccactcc 2400
 cacgatttggc gc当地atggta gaagagcagg cactgcatgt caaaaacggc ctggatgca 2460
 tccgggtctt caaggccgag cccatcggtt cactggccat cgaggaagct atggcagcat 2520
 ggtcagaaat atcagacaac ccaggacagg agcgagccac ctgcaggaa gagaaggcag 2580
 geagttcggg tctcagaaaa cc当地gcctt cagcaattgg atcaactgaa ggccgtgcac 2640
 ctgc当地atccg cggtcaggaa cctggagaga gcatgcacgc cgctgaaact ttggaaatcc 2700
 ccccaagaaa tctccaggca tcaaggactg gttacatgt ttattacgtt tatgtatcaca 2760
 ggggtgaagc ggttaaggaa atccaagatg ctgactctat catggttcaa tcaggcctt 2820
 atgggtatag caccctctca ggaggagaca atgaatctga aaacagcgtt gtggatattt 2880
 gogaacatgtt accggaggaa tatgttatca ctgaccgggg atctgtccc atctctatgg 2940
 ggttcaggcc ttctgtatgtt gaaactgcag aaggaggaaa gatccacgg ctcctgagac 3000
 tccaatccag aggcaacaac ttccgaagc ttggaaacac tctcaatgtt cctccgc当地 3060
 cggccccccg tagggccagc acttccggaa cacccattaa aaagggcaca gacggagat 3120
 tagcctcatt tggaaacggag atcgcgtttt tattgacagg tggtaaccacc caatgtgtc 3180
 gaaagtcacc ctggaaacca tcaggccag gtgcacccgc gggaaatgtc cccggatgt 3240
 tgagcaatgc cgcaatgtt caggagtgaa cacccgaaatc tggtaaccaca atctcccgaa 3300
 gatcccagaa taatgaagaa gggggagact attatgtatg tgagctgtc tctgtatgtcc 3360
 aagatattaa aacagcctt gccaaaatac acggggataa tcagaagata atctccaaac 3420
 tagaatcaatgtt gctgttattt aagggagaag ttgagtcata taagaagcag atcaacaggc 3480
 aaaatatcag catatccacc ctggaaaggac acctctcaag catcatgatc gccattctg 3540
 gacttggaa ggatccaaac gacccactg cagatgtca aatcaatccc gacttggaaac 3600
 ccatcatagg cagagattca ggccgagcac tggccgaagt tctcaagaaa cccgttgc当地 3660
 gccgacaact ccaagaaatg acaaattggac ggaccagttc cagaggacag ctgctgaagg 3720
 aatttcagct aaagccgatc gggaaaaaga tgagctcagc cgtcggtttt gttctgaca 3780
 cccggccctgc atcacgcgtt gtaatccgtt ccattataaa atccagccgg cttagggagg 3840
 atcggaaagcg ttacctgtt acttccttg atgatataa aggagccat gatcttgc当地 3900
 agttccacca gatgtgtatg aagataataa tgaagtagct acagctcaac ttacctgcca 3960
 accccatgcc agtcgacccca actagtctac cctccatcat ttttataaaa aacttaggaa 4020
 ccaggccac acagccgcca gccc当地atcaac gcgtatcttcc accgggtatc tatacgatc 4080
 ggc当地atgtt aagggagaag aacttttcac tggagttgtc coatttctt gttgatgtt 4140
 tgggtatgtt aatggggaca aattttctgtt cagtgagag ggtgaagggt atgcaacata 4200
 cggaaaaactt acccttaaat ttatgtcact tactggaaaat cttacctgtc catggccaaac 4260
 acttgc当地acttccacct atgggttca atgctttca agataccag atcatatgaa 4320
 acggcatgac ttttcaaga gtgc当地atgca cgaaggatccac gtacaggaaa gaactatatt 4380

-continued

tttcaaagat gacgggaact acaagacacg	tgctgaagtc aagtttgaag gtgataccct	4440
tgttaataga atcgagttaa aaggatttga	ttttaaagaa gatggaaaca ttcttgaca	4500
caaattggaa tacaactata actcacacaa	tgtatacatc atggcagaca aacaaaagaa	4560
tggaatcaga gttaacttca aaatttagaca	caacattgaa gatggaagcg ttcaactagc	4620
agaccattat caacaaaata ctccaattgg	cgatggccct gtcctttac cagacaacca	4680
ttacctgtcc acacaatctg cccttcgaa	agatccaaac gaaaagagag accacatggt	4740
ccttcgttag tttgtAACAG ctgctggat	tacacatggc atggatgaac tatacaaata	4800
gtgagcgcgc agcgctgacg tctcgcgatg	atactagtac aacctaatac catcataaaa	4860
aacttaggag caaagtgttatt gcctccaaag	ttccacaatg acagagatct acgacttcga	4920
caagtccggca tgggacatca aagggtcgat	cgctccgata caacccacca cctacagtga	4980
tggcaggctg gtgccccagg tcagagtcat	agatcctggt ctaggcgaca ggaaggatga	5040
atgcttatg tacatgtttc tgctgggggt	tggtgaggac agggattccc tagggctcc	5100
aatcgggcga gcatttgggt ccctgcccctt	aggtgttgc agatccacag caaagccgaa	5160
aaaactcctc aaagaggcca ctgagcttga	catagttgtt agacgtacag cagggctcaa	5220
tgaaaaactg gtgttctaca acaacacccc	actaactctc ctcacacctt ggagaaaggt	5280
cctaacaaca gggagtgtct tcaacgc当地	ccaagtgtgc aatgcggta atctgatacc	5340
gctcgatacc ccgcagaggt tccgtgttgt	ttatatgagc atcaccggc tttcgatataa	5400
cgggtattac accgttccata gaagaatgt	ggaattcaga tcggtaatg cagtgccctt	5460
caacctgctg gtgaccctta ggattgacaa	ggcgataggc cctggaaaga tcatcgacaa	5520
tacagagcaa cttcctgagg caacatttat	agtccacatc gggacttca ggagaaagaa	5580
gagtgaatgc tactctgccg attattgaa	aatgaaaatc gaaaagatgg gcctgggttt	5640
tgcacttggt gggatagggg gcaccagtct	tcacattaga agcacaggca aaatgagcaa	5700
gactctcaat gcacaactcg ggttcaagaa	gaccttatgt taccgctga tggatataa	5760
tgaagacctt aatcgattac tctggaggag	cagatgc当地 atagtaagaa tccaggcagt	5820
tttgcagcca tcagttccctc aagaattccg	catttacgac gacgtgatca taaatgatga	5880
ccaaaggacta ttcaaagttc tttttttttt	ttttttttttt agtgccc当地 aatgccc当地	5940
cttcacaatga cagccagaag gccc当地	ggccggacaa aaaagcccccc tccgaaagac tccacggacc	6000
aagcgagagg ccagccagca gccc当地	ggccggcc agcgc当地 aacgacccccc	6060
acagccctga cacaaggcca ccaccagcca	cccaatctg catcctccctc gtgggacccc	6120
cgaggaccaa ccccaaggc tgccccggat	ccaaaccacc aaccgc当地 ccaccacccc	6180
cgggaaagaa accccccagca attggaaaggc	ccctccccctt ctccctcaac acaagaactc	6240
cacaacccgaa ccgc当地 aaccgc当地	gaccgc当地 acccaaccgc aggcatccga ctccctagac	6300
agatcctctc tccccggcaa actaaacaaa	acttagggcc aaggaacata cacacccaa	6360
agaacccaga ccccgccca cggccgc当地	cccccaacc cccgacacca gagggagccc	6420
ccaaaccaatc ccgc当地 ccccggtgcc	cacaggcagg gacaccaacc cccgacacaga	6480
cccagccccc aaccatcgac aatccaagac	ggggggggccc ccccaaaaaa aagccccccag	6540
ggggccgacag ccagccaccgc gaggaagccc	acccacccca cacacgacca cggcaaccaa	6600
accagaccc accaccctt gggccaccag	ctccctccact cggccatcac cccgacagaaa	6660
ggaaaggcca caacccgc当地 accccagccc	cgatccggcg gggagccacc caacccgaac	6720
cagccacccaa gagcgatccc cgaaggaccc	ccgaacccga aaggacatca gtatcccaca	6780

-continued

gcctctccaa gtccccgggt ctccctct tctcgaaggg accaaaaat caatcccca 6840
caccggacga cactcaactc cccaccccta aaggagacac cgaaaatccc agaatcaaga 6900
ctcatccaaat gtccatcatg ggtctcaagg tgaacgtctc tgccatattc atggcagttac 6960
tgttaactct ccaaacaccc accggtaaaa tccattgggg caatctct aagatagggg 7020
tggtaggaat aggaagtgca agctacaaag ttatgactcg ttccagccat caatcattag 7080
tcataaaat aatgccccaaat ataactctcc tcaataactg cacggggta gagattgcag 7140
aatacaggag actactgaga acagtttgg aaccaattag agatgcacctt aatgcaatga 7200
cccagaatat aagaccgggtt cagagtgttag cttcaagtag gagacacaag agattgcgg 7260
gagtagtccct ggcagggtcg gcccctagcg ttgcacacg tgctcagata acggccggca 7320
ttgcacttca ccagtgccatg ctgaactctc aagccatcga caatctgaga gcgcgactgg 7380
aaactactaa tcaggcaatt gaggcaatca gacaaggagg gcaggagatg atattggctg 7440
ttcagggtgtt ccaagactac atcaataatg agctgatacc gtctatgaac caactatctt 7500
gtgatttaat cggccagaag ctcggctca aattgctcag atactataca gaaatcctgt 7560
cattattgg cccccagttta cgggacccca tatctgcggg gatataatc caggcttga 7620
gctatgcgtt tggaggagac atcaataagg tgtagaaaa gctcggatac agtggaggtg 7680
attacttggg catcttagag agcagaggaa taaaggcccc gataactcac gtcgacacag 7740
agtcctactt cattgtccctc agtatacgct atccgacgct gtccgagatt aaggggggtga 7800
ttgtccaccc gctagagggg gtctcgatca acataggctc tcaagagtgg tataccactg 7860
tgcccaagttt tgtagcaacc caagggtacc ttatctcgaa ttttgatgag tcatcgatgta 7920
ctttcatgcc agaggggact gtgtcagcc aaaatgcctt gtacccgatg agtccctctgc 7980
tccaagaatg cttccggggg tacaccaagt cctgtgctcg tacactcgta tccgggtctt 8040
ttgggaaccg gttcattttt tcacaaggga acctaatacg caatttgca tcaatccctt 8100
gcaagtgtta cacaacagga acgatcatta atcaagaccc tgacaagatc ctaacataca 8160
ttgctgcgca tcactgccc gtagtcgagg tgaacggcgt gaccatccaa gtcgggagca 8220
ggaggtatcc agacgctgt tacttgacca gaattgaccc cggtccccc atatcattgg 8280
agaggttggg cgtagggaca aatctgggg atgcaattgc taagttggag gatgccaagg 8340
aattgttggg gtcatggac cagatattga ggagttatgaa aggttatcg agcactagca 8400
tagtctacat cctgattgca gtgtgtctt gagggtttagt agggatcccc gctttatata 8460
gttgctgcag gggcgcttgtt aacaaaaagg gagaacaagt tggtatgtca agaccaggcc 8520
taaaggcctga tcttacggga acatcaaaat cctatgtaa gtcgcctctga tcccttacaa 8580
ctcttggaaac acaaattgtcc cacaagtctc ctttcgatca tcaagcaacc accgcacccca 8640
gcatcaagcc cacctgaaat tatctccggc ttccctctgg cggacaata tcggtagtta 8700
attaaaaactt agggtgcaag atcatccaca atgtcaccac aacgagaccc gataaatgcc 8760
ttctacaaatg aatccccca tcccaaggga agtaggatag tcttacacag agaacatctt 8820
atgattgata gaccttatgt tttgtggct gttctgttttgc tcatgtttct gagttgtatc 8880
gggttgcgtt ccattgcagg cattagactt catcggcag ccattcacac cgcagagatc 8940
cataaaaagcc tcagcaccaaa tcttagatgtaa actaactcaa tcgagcatca ggtcaaggac 9000
gtgtgcacac cacttccaa aatcatcggt gatgaagtgg gctgaggac acctcagaga 9060
ttctactgacc tagtggaaattt catctctgac aagattaaat tctttaatcc ggtatgggg 9120

-continued

tacgacttca gagatctcac ttggtgtatc aacccgccag agagaatcaa attggattat 9180
 gatcaatact gtgcagatgt ggctgctgaa gagctcatga atgcatttgtt gaactcaact 9240
 ctactggaga ccagaacaac caatcagttc ctagctgtct caaaggaaa ctgctcagg 9300
 cccactacaa tcagaggtca attctcaaac atgtcgctgt ccctgttaga cttgtattta 9360
 ggtcgagggtt acaatgtgtc atctatagtc actatgacat cccaggaaat gtatggggaa 9420
 acttacctag tggaaaagcc taatctgagc agcaaaaaggc cagagtgtc acaactgagc 9480
 atgtaccgag tggggggggc aatctgttatac agaaatccgg gtttgggggc tccgggttcc 9540
 catatgacaa actatcttgc gcaaccagtc agtaatgatc tcagcaactg tatgggtggct 9600
 ttggggggc tcaaactcgc agccctttgt cacggggaaag attctatcac aattccctat 9660
 caggggatcag ggaaagggtgt cagcttccag ctcgtcaagc taggtgtctg gaaatcccc 9720
 accggacatgc aatcctgggt ccccttatca acggatgatc cagtataga caggcttac 9780
 ctctcatctc acagaggtgt ttcgtctgac aatcaagcaa aatgggttgtt cccgacaaca 9840
 cgaacagatgc acaagttgcg aatggagaca tgcttccaac aggctgttaa gggtaaaatc 9900
 caagcactct gcgagaatcc cgagtgggc ccattgaagg ataacaggat tccttcatac 9960
 ggggttttgtt ctgttgatct gaggctgaca gttgagctta aaatcaaaaat tgcttcggg 10020
 ttccggccat tgatcacaca cgggttcaggg atggacctat acaaatccaa ccacaacaat 10080
 gtgttattggc tgactatccc gccaatgaag aaccttagcct taggtgttaat caacacattg 10140
 gagtggtatac cgagattcaa ggtagtccc taccttctca ctgtcccaat taaggaagca 10200
 ggcgaagact gccatgcccc aacataccta cctgcggagg tggatgggtga tgtcaactc 10260
 agttccaatc tggtgattct acctggtaa gatctccaat atgtttggc aacctacgat 10320
 acttccaggg ttgaacatgc tgggttttat tacgtttaca gcccaggccg ctcattttct 10380
 tactttatc tttttaggtt gcctataaag ggggtccccca tcgaattaca agtggaatgc 10440
 ttcacatggg accaaaaact ctggtgcctg cacttctgt tgcttgcggc ctcagaatct 10500
 ggtggacata tcactcaatc tgggtgggtg ggcattggag tcagctgcac agtcacccgg 10560
 gaagatggaa ccaatgcag atagggtgc tagtgaacca atcacaatgt gtcacccaga 10620
 catcaggcat acccactagt gtgaaataga catcagaatt aaaaaaaaaacg tagggtccaa 10680
 gtgggtcccccc gttatggact cgctatctgtt caaccagatc ttataccctg aagttcacct 10740
 agatagccccg atagttacca ataagatagt agccatctg gaggatgtc gaggccctca 10800
 cgcttacagc ctggaggacc ctacactgtc tcagaacatc aagcacccgg taaaaaacgg 10860
 atttccaaac caaatgatta taaacaatgt ggaagttggg aatgtcatca agtccaaagct 10920
 taggagttat ccggccact ctcataatcc atatccaaat tgtaatcagg atttattaa 10980
 catagaagac aaagagtcaa cgaggaagat ccgtgaactc ctcaaaaagg ggaattcgct 11040
 gtactccaaa gtcagtgata aggtttcca atgcttaagg gacactaact cacggcttgg 11100
 cctaggctcc gaatttgagg aggacatcaa ggagaaagt attaacttgg gagttacat 11160
 gcacagctcc cagtggttttgc agccctttctt gtttgggtt acagtcaaga ctgagatgag 11220
 gtcagtgattt aaatcacaaa cccatacttg ccataggagg agacacacac ctgtatttt 11280
 cactggtagt tcagttgagt tgctaatctc tcgtgacattt gttgtataa tcagtaaaga 11340
 gtctcaacat gtatattacc tgacatttgc actgggttttgc atgtatttgtt atgtcataga 11400
 ggggaggtta atgacagaga ccgtatgac tattgtatgc aggtatacag agcttctagg 11460
 aagagtcaga tacatgtgga aactgtataga tggtttcttc cctgcactcg ggaatccaa 11520

-continued

ttatcaaatt gtagcaatgc tggagccctt ttcacttgct tacctgcagc tgagggatat 11580
 aacagtagaa ctcagaggtg ctcccttaa ccactgctt actgaaatac atgatgttct 11640
 tgacccaaaac gggtttctg atgaaggta tacatcatgag ttaattgaag ctctagatta 11700
 cattttcata actgatgaca tacatctgac aggggagatt ttctcattt tcagaagttt 11760
 cggccacccc agacttgaag cagtaacggc tgctgaaaat gttaggaat acatgaatca 11820
 gectaaagtc attgtgtatg agactctgat gaaaggtcat gccatattt gtggaatcat 11880
 aatcaacggc tategtgaca ggcacggagg cagttggcca ccgctgaccc tccccctgca 11940
 tgctgcagac acaatccgga atgctcaagc ttcaaggta gggtaaacac atgagcagt 12000
 cgttgataac tggaaatctt ttgctggagt gaaatttggc tgctttatgc ctcttagcct 12060
 ggatagtgtatg ctgacaatgt acctaaagga caaggcactt gctgctctcc aaagggatg 12120
 ggattcagtt taccggaaag agttcctgcg ttacgaccct cccaaaggaa ccgggtcacg 12180
 gaggcttgcgatgtttcc ttaatgatcc gagcttgcg ccataatgt tgataatgt 12240
 tggtaagt ggagcttacc tccatgaccc tgagttcaac ctgtcttaca gcctgaaaga 12300
 aaaggagatc aaggaaacag gttagacttt tgctaaaatg acttacaaa tgagggcatg 12360
 ccaagtgtatt gctgaaaatc taatctcaa cgggattggc aaatattta aggacaatgg 12420
 gatggccaaag gatgagcactg atttgactaa ggcactccac actcttagctg tctcaggagt 12480
 ccccaaagat ctcaaaagaaa gtcacagggg gggccagtc ttaaaaacct actcccgaaag 12540
 cccagtcac acaagttacca ggaacgttag agcagcaaaa gggtttatag gttccctca 12600
 agtaattcgg caggaccaag acactgatca tccggagaat atggaagctt acgagacagt 12660
 cagtgcattt atcacgactg atctcaagaa gtactgcctt aattggagat atgagaccat 12720
 cagcttgcattt gcacagaggg taaatgagat ttacggattt ccctcatttt tccagtggct 12780
 gcataagagg cttgagacct ctgtcctgta tgtaagtgc cctcattgcc ccccgacact 12840
 tgacgccccat atcccgatataaaggatccc caatgatcaa atcttcattt agtaccctat 12900
 gggaggatataaaggatccc gtcagaagct gtggaccatc agcaccatc cctatctata 12960
 cctggctgct tatgagagcg gagtaaggat tgcttcgtta gtgcaaggaa acaatcagac 13020
 catagccgtaa caaaaagggtt taccggacac atggccctac aacctaaga aacgggaagc 13080
 tgcttagatca acttagatca actttgtatc tcttaggcaaa aggctacatg atattggcca 13140
 tcacccatca gcaaatgaga caatttttcc atcacattttt tttgtcttattt caaaaggat 13200
 atattatgtt gggctacttg tgtcccaatc actcaagagc atcgcaagat gtgtattctg 13260
 gtcagagact atagttgtatc aaacaagggc agcatgcagt aatattgcta caacaatggc 13320
 taaaaggatc gagagaggtt atgaccgtta ccttgcataatccctgaaacg tcctaaaatgt 13380
 gatacagcaa attctgtatct ctcttggctt cacaatcaat tcaaccatga cccggatgt 13440
 agtcataccccc ctcctcacaatc acaacgacactt cttataagg atggcactgt tgccgcctcc 13500
 tattgggggg atgaattatc tgaatatgag caggctgtttt gtcagaaaca tcggtgatcc 13560
 agtaacatca tcaattgctg atctcaagag aatgatttcc gctctactaa tgcctgaaga 13620
 gaccctccat caagtaatga cacaacaacc gggggactct tcattccatg actgggctag 13680
 cgacccttac tcaatgttcc ttccttgcattt ccaatccatgacttcc tcaagaacat 13740
 aactgcaagg tttgtccatgatc tccatagttcc aaacccaaatg ttaaaaaggat tattccatga 13800
 tgacagtaaaa gaagaggacg agggactggc ggcattccctc atggacagggc atattatgt 13860

-continued

acctaggcca gctcatgaaa tcctggatca tagtgtcaca gggcaagag agtctattgc	13920
aggcatgctg gataccacaa aaggcttgcat tcgagccagc atgaggaagg gggggtaac	13980
ctctcgagtg ataaccagat tgtccaatta tgactatgaa caattcagag cagggatgtt	14040
gctattgaca ggaagaaga gaaatgtcct cattgacaaa gagtcatgtt cagtgcagct	14100
ggcgagagct ctaagaagcc atatgtgggc gaggctagct cgaggacggc ctatsscgg	14160
ccttgaggc cctgatgtac tagaatctat gegaggecac cttattcggc gtcatgagac	14220
atgtgtcatac tgcgagtgtg gatcagtcaa ctacggatgg tttttgtcc cctcggttg	14280
ccaaactggat gatattgaca agggaaacatc atccttgcgat gtcggatata ttgggttctac	14340
cactgatgag agaacagaca tgaagcttgc cttcgtaa gccccaaatgc gatccttgcg	14400
atctgctgtt agaatagcaa cagtgtactc atgggcttac ggtgatgtat atagcttttg	14460
gaacgaagcc tgggtgttgg ctaggcaag ggccaatgtg agcctggagg agctaagggt	14520
gatcactccc atctcaactt cgactaattt agcgcataagg ttgagggatc gtacactca	14580
agtgaaatac tcaaggatcat cccttgcgg agtggcgagg tataccacaa tctccaaacga	14640
caatctctca tttgtcatat cagataagaa ggttataact aacttttat accaacaagg	14700
aatgctccta ggggtgggtt ttttagaaac attgtttcga ctcgagaaag ataccggatc	14760
atctaacaacg gtattacatc ttacgtcga aacagattgt tgcgtgatcc cgatgtataga	14820
tcatcccagg atacccagct cccgcaagct agagctgagg gcagagctat gtaccaaccc	14880
attgatatat gataatgcac cttaattga cagagatgca acaaggctat acacccagag	14940
ccataggagg caccttgcgg aatttgcgtt atgggtccaca ccccaactat atcacattt	15000
agctaagtcc acagcactat ctatgattga cctggtaaca aaatttgcgaa aggaccatata	15060
gaatgaaatt tcaagctctca tagggatga cgatataat agtttgcataa ctgagttct	15120
gctcatagag ccaagattat tcaactatcta cttggggccag tgcgtggccca tcaattggc	15180
atttgcgtt cattatcata gaccatcagg gaaatatcgtt atgggtgagc tgggtcatac	15240
gttccttttctt agaatgagca aaggagtgtt taagggtgtt gtcaatgctc taagccaccc	15300
aaagatctac aagaatttctt ggcattgtgg tattatagag cctatccatg gtccttca	15360
tgtatgcataa aacttgcaca caactgtgtt caacatgggt tacacatgtt atatgacta	15420
cctcgacccgtg ttgttgcata aagagtttgcg agagttcaca ttatcttgcgtt gtgaaagcga	15480
cgaggatgtt gtaaccggaca gattcgacaa catccaggca aaacacttat gtgttgcgc	15540
agatttgcgtt tgcgttgcaca ggacccgttcc accaattcgaa ggtctaaagac cggtagagaa	15600
atgtgcgtt ctaaccggacc atatcaaggc agaggctatg ttatctccatc caggatctt	15660
gtggaaacata aatccaattt ttgttagacca ttactcatgc tctctgactt atctccggcg	15720
aggatcgatc aaacagataa gattgagatg tgcgttgcgc ttcatatccgc acgcctcgc	15780
tgaggtaaat gtcgttgcgc caaagatcgatc cagcaacaatc atctcaataa tgagcatcaa	15840
ggctttcaga ccccccacacg atgatgttgc aaaatttgcgc aaagatatac acacaagca	15900
gcacaatctt cccatttcag gggcaatctt cgccaaatattt gaaatccatg cttccgcag	15960
aatcggttttgc aactcatctt cttgttgcataa agctgttgcgtt atatcaacat taatttgcgtt	16020
atgccttgcgtt ccaggggagg acggcttgcgtt cttgggtgcgtt ggatcggttgcgtt cttatgttgcgtt	16080
cacttataag gagataactta aactaaacaa gtgttgcataa aatgtgggg tttccggccaa	16140
ttcttagatctt ggtcaaaaggg aatttgcacc ctatccctcc gaaatggggcc ttgttgcataa	16200
cagaatggga gtaggtataa ttgttgcataa gtcctttaac gggaggcccg aagtcacgtt	16260

-continued

ggttaggcagt gtagattgct tcaatttcat agttagtaat atccctacact ctatgtggg 16320
 gtttatccat tcagatatacg agacccgtcc tgacaaagat actatagaga agcttagagga 16380
 attggcagcc atcttatcga tggctctgtc cctggggaaa ataggatcaa tactggtgat 16440
 taagcttatg ctttcagcg gggatttgt tcagggattt ataagttatg tagggctca 16500
 ttatagagaa gtgaaccttg tataccctag atacagcaac ttcatatcta ctgaatctta 16560
 ttggttatg acagatctca aggctaaaccg gctaataaatc cctgaaaaga ttaaggcagca 16620
 gataattgaa tcatctgtga ggacttcacc tggacttata ggtcacatcc tatccattaa 16680
 gcaactaagg tcgcatacaag caattgtggg agacgcaggat agtagaggtg atatcaatcc 16740
 tactctgaaa aaacttacac ctatagagca ggtgctgatc aattgcgggt tggcaattaa 16800
 cggacctaag ctgtgcaaaag aattgtatcca ccatgtatgtt gcctcagggc aagatggatt 16860
 gettaattct atactcatcc tctacaggga gttggcaaga ttcaaaagaca accaaagaag 16920
 tcaacaaggg atgttccacg cttacccgtt attggtaagt agcaggcaac gagaacttat 16980
 atcttaggatc accccaaat tttggggca catttttttt tactccggg acagaaagtt 17040
 gataaataag tttatccaga atctcaagtc cggttatctg atactagact tacaccagaa 17100
 tatcttcgtt aagaatctat ccaagtccaga gaaacagattt attatgacgg ggggtttgaa 17160
 acgtgagtgg gtttttaagg taacagtccaa ggagacccaa gaatggtata agttagtcgg 17220
 atacagtgcc ctgattaagg actaatttgtt tgaactccgg aaccctaaatc ctgccttagg 17280
 tggtaggca ttatggcaaa tataattaaag aaaacttgcgaa aataacgaaat tttctattcc 17340
 cagcttgc tggtagggccgg catggtccca gcctccctcg tggcgcggc tggcaacat 17400
 tccgaggggg ccgtcccccgg ggtatggcg aatgggacccg ggccgatccg gctgtaaca 17460
 aagcccgaaa ggaagctgag ttggctgctg ccaccgctga gcaataacta gcataacccc 17520
 ttggggccctc taaaagggtc ttgaggggtt tttgctgaa aggaggaact atatccggat 17580
 gccggcccgag gtaccggcgtt ttagtggggg ttaatttgcgat gcttggcgta 17640
 atcatggtca tagctgttccct ctgtgtgaaa ttgttatccg ctcacaatcc cacacaacat 17700
 acgagccgga agcataaaagt gtaaaggctcg ggggtgcctaa tgagtggact aactcacatt 17760
 aattgeggtt cgctactgc ccgtttccaa gtcggggaaac ctgtcgatcg agctgcattta 17820
 atgaatcgcc caacgcgcgg ggagagggcg tttgcgtattt gggcgctttt ccgttccctc 17880
 gctcaactgac tggctgtcgat cggctgtcgat gctgcccgcg ggggtatccg ctcactcaaa 17940
 ggccggtataa cgggtatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 18000
 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcggttt tccataggct 18060
 ccggcccccgtt gacgagccatc acaaaaatcg acgctcaagt cagagggtggc gaaacccgac 18120
 agactataaa agataaccagg cgtttccccc tggaaagctcc ctcgtgcgtt ctcctgttcc 18180
 gaccctgcggc cttaccggat acctgtccgc ctttctccct tcggaaagcg tggcggtttc 18240
 tcatacgatca cgctgttaggt atctcagttc ggtgttaggtc gttcgatccaa agctgggtcg 18300
 tggcactgaa ccccccgttc agcccgacccg ctgcgcctta tcggtaact atcgatctga 18360
 gtccaaaccgg gtaagacacg acttacgcgc actggcagca gccactggta acaggattag 18420
 cagagccgagg tatgttagggcg gtgcgtacaga gttcttgaag tggtagggccata actacggctaa 18480
 cactagaagg acagtatccgtt gatatctgcgc tctgctgaaag ccagttaccc tcggaaaaag 18540
 agttggtagc tcttgcgttccg gcaaaacaaac caccgctggt agcgggtggtt tttttgttt 18600

-continued

```

caaggcagcag attacgcgca gaaaaaaaagg atctcaagaa gatccttga tctttctac 18660
gggggtctgac gtcagtggta acgaaaactc acgttaaggg atttggtca tgagattatc 18720
aaaaaggatc ttcacctaga tcctttaaa taaaaatga agtttaat caatctaaag 18780
tatatatatgag taaacttgtt ctgacagtta ccaatgttta atcagtgagg cacctatctc 18840
agcgatctgt ctatccgtt catccatagt tgctgtact cccgtctgt agataactac 18900
gatacgggg ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccacgctc 18960
accggctcca gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg 19020
tcctgcaact ttatccgcct ccattccgtc tattatgt tgccggaaag cttagtaag 19080
tagttcgcca gttaatagtt tgcgcaacgt tggtgccatt gctacaggca tcgtggtgc 19140
acgctcgctg tttgttatgg cttcattcag ctccgggtcc caacgatcaa ggcgagttac 19200
atgatcccc atgttgtgca aaaaagcggt tagtccttc ggtcctccga tcgttgtcag 19260
aagtaagttg gccgcagtg tatcactcat ggttatggca gcactgcata attctcttac 19320
tgcatgcca tccgtaagat gctttctgt gactggtag tactcaacca agtcattctg 19380
agaatagtgt atgcggcgac cgagttgtc ttgcccggc tcaatacggg ataataccgc 19440
gccacatago agaactttaa aagtgcctat cattggaaaa cgttcttcgg ggcgaaaact 19500
ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgta caccaactg 19560
atcttcagca tctttactt tcaccagctg ttctgggtga gaaaaaacag gaaggaaaa 19620
tgccgcaaaa aagggataa gggcacacg gaaatgttg atactcatac ttttccttt 19680
tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg 19740
tatttagaaa aataaacaaa taggggttcc gcgcacattt cccgaaaaag tgc 19793

```

<210> SEQ ID NO 2
<211> LENGTH: 19798
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Complete nucleotide sequence of p(+)MV3EZ-GFP

<400> SEQUENCE: 2

```

cacctaaatt gtaagcgtta atattttgtt aaaattcgcg ttaaattttt gttaaatcag 60
ctcattttt aaccaatagg ccgaaatcg 3' caaaatccct tataaatcaa aagaatagac 120
cgagataggg ttgagtgttg ttccagtttgc 3' gaacaagagt ccactattaa agaacgtgga 180
ctccaaacgtc aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc 240
accctaatac agtttttgg ggtcgagggtg ccgtaaagca ctaaatcgga accctaaagg 300
gagcccccgta tttagagctt gacggggaaa gccggccatt taggccccatggc 360
agttagcggtt ctaacgtgtcg 3' cgtaaccacc acacccgcgg cgcttaatgc gcccgtac 420
ggcgcgatccc attcgcatt caggctgcgc 3' aactgttggg aaggggcgatc ggtggggcc 480
tcttcgtat taacgttgcgtt ggcgaaagggg ggatgtgtcg 3' caaggcgatt aagttggta 540
acggccagggtt tttccagtc acgacgttgtt 3' aaaacgcacg ccagtgaatt gtaatacgc 600
tcactataac caaacaaagt tgggttaagga tagttcaatc aatgtatcattt ttcttagtgc 660
cttaggattc aagatcctat tatacggac 3' aagagcagga ttagggatat ctgagatggc 720
cacactttta aggagcttag cattgttcaa 3' aagaaacaag gacaaaccac ccattacatc 780
aggatccggtt ggagccatca gaggaatcaa acacattttt atagttccaa tccctggaga 840
ttcctcaattt accactcgat ccagacttctt ggaccgggttgc 3' gtcaggtaa ttggaaaccc 900

```

-continued

ggatgtgagc gggcccaaac taacaggggc actaataggta atattatcct tatttggaa	960
gtctccaggtaaatttgcattc agaggatcac cgatgaccct gacgtagca taaggctgttt	1020
agagggtgtc cagagtgacc agtcacaatc tggccttacc ttgcataa gaggtaccaa	1080
catggaggat gaggcggacc aataacttttc acatgtatcatc ccaatttagta gtgtatcaatc	1140
cagggttcgga tggttcgaga acaaggaaat ctcagatattt gaagtgcag accctgagg	1200
attcaacatcg attctgggta ccattccttcg ccaaatttgg gtcttgctcg caaaggcggt	1260
tacggcccca gacacggcag ctgattcgga gctaagaagg tggataaaatg acacccaaca	1320
aagaagggtt gtttgtaat tttagttggaa gagaatggg ttggatgtgg tgaggaaacag	1380
gattgcggag gaccccttcct tacggccgatt catggtcgct ctaatcctgg atatcaagag	1440
aacaccccgaa aacaaacccca ggattgctga aatgtatgtt gacattgtatacatatatcg	1500
agaggcagga tttagccagtt ttatcctgcac tattaatgtt gggatagaaa ctatgtatcc	1560
tgctcttggaa ctgcataat ttgctgtgtt gttatccaca cttgagtcct tggatgttgc	1620
ttaccagcaa atgggggaaa ctgcacccta catggtaatc ctggagaact caattcagaa	1680
caagttcagt gcaggatcat accctctgtt ctggagatctt gocatgggag tagggatgtt	1740
acttgaaaac tccatgggggg gtttgaactt tggccgatctt tactttgtatc cagcatat	1800
tagattagggg caagagatgg taaggaggc agctggaaatg gtcagttccaa cattggcattc	1860
tgaactcggt atcaactgccc aggatgtcaag gcttggatca gagattgcaaa tgcataactac	1920
tgaggacaag atcagtagag cggttggacc cagacaagcc caagatcat ttctacacgg	1980
tgatcaaagt gagaatgagc taccgagattt gggggggcaag gaagatagga gggtaaaca	2040
gagtcgagga gaagccagg agagctacag agaaacccggg cccagcagag caagtgtatgc	2100
gagagctgccat cttccaa ccggcacacc cctagacattt gacactgcattt cggagtccag	2160
ccaagatcccg caggacagtc gaaggcgc tgcgccttgc cttaggctgc aagccatggc	2220
agaatctcg gaagaacaag gctcagacac ggacacccctt atagtgatca atgacagaaa	2280
tcttcctagac taggtgcgag agggcgagg ccagaacaatccgccttac cctccatcat	2340
tgttataaaa aacttaggaa ccaggtccac acagccgcca gcccatcaac catccactcc	2400
cacgattggaa gccaatggta gaagagcagg cacgcatgtt caaaaacggta ctggatgtca	2460
tccgggtctt caaggcccgag cccatcggtt cactggccat cgaggaatgtt atggcagcat	2520
ggtcagaaat atcaagacaac ccaggacagg agcgagccat ctgcaggaa gagaaggcag	2580
gcagttcggg tctcagaaaa ccatgcctt cagcaattgg atcaactgaa ggcgggtgcac	2640
ctcgcatcccg cggtcaggaa cctggagaga gcatgtacgcg cgtgtaaactt ttggatgtt	2700
ccccaaatggaa tctccaggca tcaagcactg ggttacatgtt ttattacgtt tatgtatca	2760
gcgggtgaagc ggtttaaggaa atccaagatgtt ctgactctat catgggttca tcaggccctt	2820
atgggtatgtt caccctctca ggaggagaca atgaatgtt aaacagcgat gtggatattt	2880
gcgaacctgtt taccgaggaa tatgtatca ctgaccgggg atctgttccc atctctatgg	2940
ggttcaggcc ttctgtatgtt gaaactgcag aaggaggggaa gatccacggatccttgcag	3000
tccaatccag aggcaacaac ttccgcagc ttggaaaac tctcaatgtt cctccgc	3060
cgccggccggg tagggccagc acttccggaa caccattaa aaagggcaca gacgcgagat	3120
tagccttattt tggaaacggag atcgcgtttt tatttgcagg tgggtgcaccatgtgtc	3180
gaaaatgcaccatc tcggaaacca tcaggccag gtcacccgtc gggaaatgtc cccgagtg	3240

-continued

ttagcaatgc	cgcactgata	caggagtgg	cacccgaatc	tggtaccaca	atctccccga	3300
gatcccagaa	taatgaagaa	gggggagact	attatgtat	tgagctgttc	tctgtatgtcc	3360
aagatattaa	aacagcctt	gccaaaatac	acgaggataa	tcagaagata	atctccaagc	3420
tagaatact	gctgttattt	aagggagaag	ttgagtcaat	taagaagcag	atcaacaggc	3480
aaaatatcg	cataatccacc	ctggaaggac	acctctcaag	catcatgatc	gccattctg	3540
gacttggaa	ggatccaaac	gaccccaact	cagatgtcg	aatcaatccc	gacttgaac	3600
ccatcatagg	cagagattca	ggccgagcac	tggccgaagt	tctcaagaaa	cccggttgc	3660
gcccacaact	ccaaggaaat	acaaatggac	ggaccagttc	cagaggacag	ctgctgaagg	3720
aatttcagct	aaagccgatc	ggaaaaaaga	tgagctcagc	cgtcggtt	gttccgtaca	3780
ccggccctgc	atcacgcgt	gtaatccgt	ccattataaa	atccagccgg	ctagaggagg	3840
atcggaaagcg	ttacctgat	acttcctt	atgatataa	aggagccaat	gatcttgc	3900
agttccacca	gatgtgtat	aagataataa	tgaagttagct	acagctcaac	ttacctgc	3960
accccatgcc	agtcgaccca	actagtacaa	cctaaatcca	tcataaaaaa	cttaggagca	4020
aagtgattgc	ctcccaagtt	ccacaatgac	agagatctac	gacttcgaca	agtccggat	4080
ggacatcaaa	gggtcgatcg	ctccgataca	acccaccacc	tacagtgtat	gcaggcttgt	4140
gccccagg	tcagtcatag	atccctggtct	aggcgacagg	aaggatgaat	gctttatgt	4200
catgtttctg	ctgggggtt	ttgaggacag	ggattcccta	gggcctccaa	tcggcgagc	4260
atttgggtcc	ctgcccctt	gtgttggcag	atccacagca	aagccgaaa	aactcctcaa	4320
agaggccact	gagcttgaca	tagttgtt	acgtacagca	gggctcaatg	aaaaacttgt	4380
gttctacaac	aacacccac	taactctcct	cacaccttgg	agaaagg	tccaaacagg	4440
gagtgttcc	aacgc当地	aagtgtgca	tgcggtaat	ctgataccgc	tcgataaccc	4500
gcagagg	ttc	ctgttgcgtt	atatgagcat	cacccgtt	tcggataacg	4560
cgttcc	ctaga	agaatgtcg	aattcagatc	ggtcaatgca	gtggccttca	4620
gaccctt	agg	attgacaagg	cgataggccc	tggaaagatc	atcgacaata	4680
tcctgagg	ca	acatttatag	tccacatgg	gaacttcagg	agaaagaaga	4740
ctctgccc	gat	tattgcaaaa	tgaaaatcg	aaagatggc	ctggtttt	4800
gatagggggc	acc	acttc	acattagaag	cacaggcaaa	atgagaaga	4860
acaactcggg	ttcaagaaga	ccttatgtt	cccgctgtat	gatataatg	aagacctt	4920
tcgattactc	tggaggagca	gatgtat	agtaagaatc	caggcgtt	tgcagccatc	4980
agttcc	caa	gaattccgca	tttacgacg	cgtgatcata	aatgtatgacc	5040
caaagttctg	tagaccgt	tgcccagca	tgcggaaa	cgacccccc	cacaatgaca	5100
gcacagaaggc	ccggacaaa	aagccccc	cgaaagactc	cacggacaa	gcgagaggcc	5160
agccagc	gcacggcaag	cgc当地	aggcgcccc	agcacacaac	agccctgaca	5220
caaggccacc	acc	ccatctgca	tcctcctcg	gggaccc	aggaccaacc	5280
cccaagg	ctcc	cccccate	aaaccacca	ccgc当地	accaccc	5340
ccccagcaat	tggaaagg	cccccct	tcctcaacac	aagaactcc	caaccgaacc	5400
gcacaagc	ccgagg	tgac	ccaaaccgc	gcatcc	ccctagacag	5460
ccggcc	caa	taaacaaac	ttagggccaa	ggaacataca	caccaacag	5520
ccggcc	acg	ccggccgccc	cccaacccc	gacaacc	gggagcccc	5580
gcgg	ctccc	ccgggtcccc	caggcagg	caccaacccc	cgaacagacc	5640

-continued

ccatcgacaa tccaagacgg gggggcccc caaaaaaaaa gcccccaggg gccgacagcc	5700
agcacccgca ggaagccac ccacccaca cacgaccacg gcaaccaaac cagaaccag	5760
accacccctgg gccaccagct cccagactcg gccatcaccc cgccagaaagg aaaggccaca	5820
acccgcgcac cccagccccg atccggcggg gagccaccca acccgaacca gcacccaaga	5880
gегатccccg aaggacccc gaaccgcaaa ggacatcagt atcccacagc ctctccaagt	5940
ccccccgtct ctcctcttc tcgaaggac caaaagatca atccaccaca cccgacgaca	6000
ctcaactccc cacccctaaa ggagacaccc ggaatcccag aatcaagact catccaatgt	6060
ccatcatggg tctcaagggtg aacgtctcg ccatattcat ggcaactctg ttaactctcc	6120
aaacacccac cggtcaaatac cattggggca atctctctaa gataggggtg gtaggaatag	6180
gaagtgcag otacaaagtt atgactcggt ccagccatca atcattatgc ataaaattaa	6240
tgc当地at aactctctc aataactgca cgagggtaga gattgcagaa tacaggagac	6300
tactgagaac agttttggaa ccaatttagag atgcacttta tgcaatgacc cagaatataa	6360
gaccgggtca gagtgtagct tcaagtagga gacacaagag aagttgcgaa tggagacatg	6420
cttccaaacg gcgtgttaagg gtaaaatcca agcaactctgc gagaatcccg agtgggcacc	6480
attgaaggat aacaggattc cttcatacgg ggtttgtct gttgatctga gtctgacagt	6540
ttagctaaa atcaaaattt cttcgggatt cgggcccattt atcacacacg gttcaggat	6600
ggacctatac aaatccaacc acaacaatgt gtattggctg actateccgc caatgaagaa	6660
cctagectta ggtgtaatca acacatttgc gtggataccg agatcaagg ttgtcccta	6720
cctcttcaact gtcccaatta aggaagcagg cgaagactgc catgccccaa catactacc	6780
tgccggaggtg gatggtgatg tcaactctcg ttccaatctcg gtgattctac ctggtaaga	6840
tctccaaat gtttggcaa cctacgatac ttccagggtt gaacatgctg tggtttatta	6900
cgtttacagc ccaggccgct cattttctta cttttatctt ttaggttgc ctataaagg	6960
ggtccccatc gaattacaag tggatgtt cacatggac caaaaactct ggtgcgtca	7020
cttctgtgtg cttgeggact cagaatctgg tggacatatc actcaactctg ggtgggtgg	7080
catgggagtc agctgcacag tcacccggga agatggaacc aatcgacat agggtctgta	7140
gtgaaccaat cacatgtatgt cacccagaca tcaggcatac ccactagtc accctccatc	7200
attgttataa aaaacttagg aaccaggatcc acacagccgc cagcccatca acgcgtatct	7260
tcacccgtga tctatcgct acgtacgtcg catgatggaa ggagaagaac tttctactgg	7320
agttgtccca attcttggat aattagatgg ttagtgcataat gggcacaat tttctgtcag	7380
tggagaggggt gaagggtatg atttgcggga gttagtgcgtt caggtgcggc cctaggcggt	7440
gccacagctg ctcagataac ggccggcatt gcacttcacc agtccatgtc gaactctcaa	7500
gccccatcaca atctgagagc gagcctggaa actactaatac aggcaattga ggcaatcaga	7560
caagcaggc aggagatgtt atttgcgtt caggtgtcc aagactacat caataatgag	7620
ctgataaccgt ctatgaacca actatcttgt gatataatcg gccagaagct cgggtcaaa	7680
ttgctcagat actatacaga aatctgtca ttatggcc coagttacg ggaccacata	7740
tctgcggaga tatctatcca ggctttgacg tatgcgtt gaggagacat caataaggt	7800
tttagaaaagc tcggatacag tggaggtatg ttactggca tcttagagag cagaggaata	7860
aaggccccggtaa taactcacgt cgacacagag tcctacttca ttgtccctcag tatacgctat	7920
ccgacgctgt ccgagattaa ggggggtgatt gtccaccggc tagaggggtt ctcgtacaac	7980

-continued

ataggctctc aagagtggta taccactgtg cccaagtatg ttgcaaccca agggtaccc	8040
atctcgaatt ttgatgagtc atcggtgtact ttcatgcggc agggggactgt gtgcagccaa	8100
aatgccttgt acccgatgag tcctctgtc caagaatgcc tccgggggta caccaagtcc	8160
tgtgctcgta cactcgatc cgggtctttt gggAACGGT tcattttatc acaaggaaac	8220
ctaataggcca atttgccatc aatcctttgc aagtgttaca caacaggaaac gatcattaat	8280
caagaccctg acaagatcct aacatacatt gctgcccgtc actgccccgt agtcgaggtg	8340
aacggcgtga ccatccaagt cgggagcagg aggtatccag acgctgtgta cttgcacaga	8400
attgaccccg gtcctccat atcattggag aggttggacg tagggacaaa tctggggaaat	8460
gcaattgcta agttggagga tgccaaggaa ttgttggagt catcgacca gatattgagg	8520
agtatgaaag gtttatcgag cactagcata gtctacatcc tgattgcagt gtgtcttgg	8580
gggtttagatgg gatccccgc ttatatatgt tgctgcaggg ggcgttgtaa caaaaaggaa	8640
gaacaagttg gtatgtcaag accaggccta aagcctgatc ttacggaaac atcaaataatcc	8700
tatgtaaagg ctgtctgtatc ctctacaact cttgaaacac aaatgtccca caagtctcc	8760
cttcgtcatc aagcaaccac cgccaccccgatc atcaagccca cctgaaatata tctccggctt	8820
ccctctggcc gaacaatatac ggtatgttataaacttag ggtgcaagat catccacaat	8880
gtcaccacaa cgagaccgga taaatgcctt ctacaaagat aaccccccata ccaaggaaag	8940
taggatagtc attaacagag aacatcttat gattgataga ctttatgttt tgctggctgt	9000
tctgtttgtc atgtttctga gcttgatcgg gttgcttagc attgcaggca ttagacttca	9060
tcgggcagcc atctacaccg cagagatcca taaaagcctc agcacaatc tagatgtaac	9120
taactcaatc gagcatcagg tcaaggacgt gctgacacca ctcttcaaaa tcatcggtga	9180
tgaagtgggc ctgaggacac ctcagagatt cactgaccta gtgaaattca tctctgacaa	9240
gattaaattc cttatccgg atagggagta cgacttcaga gatctcaccc ggtgtatcaa	9300
cccgccagag agaatcaaata tggattatga tcaatactgt gcagatgtgg ctgctgaaga	9360
gctcatgaat gcattggta actcaactct actggagacc agaacaacca atcagttcc	9420
agctgtctca aaggaaact gctcaggcc cactacaatc agaggtcaat tctcaaacat	9480
gtcgctgtcc ctgttagact tggatgtttagc tcgagggtac aatgtgtcat ctatagtc	9540
tatgacatcc caggaaatgt atggggaaac ttacctatgt gaaaagccta atctgagcag	9600
caaaaggta gagttgtcac aactgagcat gtaccgagtg tttgaagtagt gtgttac	9660
aaatccgggt ttgggggctc cgggtttcca tatgacaaac tatcttgagc aaccagtcc	9720
taatgatctc agcaactgtt tgggtggctt gggggagetc aaactcgccag ccctttgtca	9780
cggggagat tctatcacaa ttccctatca gggatcaggaa agaggtgtca gcttccagct	9840
cgtcaagcta ggtgtctggaa atccccaaac cgacatgcaatc tccctgggtcc ccttata	9900
ggatgatcca gtatgatgaca ggcttacatc ctatctcactc agaggtgtta tcgctgacaa	9960
tcaagcaaaa tgggtgtcc cgacaacacg aacagatgac caacatacgg aaaacttacc	10020
ctttaatttta ttgtcactac tggaaaacta cctgttccat gccaaacact tgcactact	10080
ttcacctatg gtgttcaatg ctttcaaga tacccagatc atatgaaacg gcatgactt	10140
ttcaagagtg ccatgcccga aggttacgtt cagggaaagaa ctatattttt caaagatgac	10200
ggaaactaca agacacgtgc tgaagtcaag ttgttggatgtt atacccttgtt taatagaatc	10260
gagttaaaag gtattgattt taaaaggat ggaaacattt ttggacacaa attgaaatac	10320
aactataact cacacaatgt atacatcatg gcagacaaac aaaagaatgg aatcagagtt	10380

-continued

aacttcaaaa ttagacacaa cattgaagat ggaagcgttc aactagcaga ccattatcaa 10440
 caaaaatactc caattggcga tggccctgtc ctttaccag acaaccatca cctgtccaca 10500
 caatctgccc ttccgaaaga tcccaacgaa aagagagacc acatggcct tcttgagtt 10560
 gtaacagctg ctgggattac acatggcattg gatgaactat acaaataatg agcgccgc 10620
 getgacgtct cgcatgata ctatgtgaa atagacatca gaattaagaa aaacgttagg 10680
 tcctcaatgtt tccccgttat ggactcgcta tctgtcaacc agatcttata ccctgaagtt 10740
 cacctagata gcccgtatgt taccaataag atagtagcca tcctggagta tgctcgagtc 10800
 cctcagcgtt acagectgga ggaccctaca ctgtgtcaga acatcaagca ccgcctaaaa 10860
 aacggatttt ccaaccaaat gattataaac aatgtgaaag ttggaaatgt catcaagtcc 10920
 aagctttagga gttatccggc ccactctcat attccatatac caaattgtaa tcaggattta 10980
 tttaacatag aagacaaaga gtcaacgagg aagatccgtg aactcctcaa aaagggaaat 11040
 tcgctgtact ccaaagtcaag tgataaggtt ttccaatgtc taagggacac taactcacgg 11100
 ctggccttag gtcggaaatt gagggaggac atcaaggaga aagttattaa cttggagtt 11160
 tacatgcaca gtcggcgtg gtttgcggc ttctgtttt gtttacagt caagactgag 11220
 atgaggtcag tgattaaatc acaaaccat acttgcataa ggaggagaca cacacctgta 11280
 ttcttcactg gtagttcgt tgatgtgta atctctcgat accttgcgtc tataatcagt 11340
 aaagagtctc aacatgtata ttacctgaca tttgaactgg ttttgcgttgc ttgtgtgtc 11400
 atagagggaa ggttaatgac agagaccgt atgactattt atgcttagta tacagagtt 11460
 ctaggaagag tcaagatacat gtggaaactg atagatggtt ttctccctgc actcggaaat 11520
 ccaacttatac aaattgttagc aatgtggag cctcttcac ttgttacact gcagctgagg 11580
 gatataacag tagaactcag aggtgcattt cttaccact gtttactga aatacatgat 11640
 gttcttgacc aaaacgggtt ttctgtgaa ggtacttatac atgagttat tgaagctcta 11700
 gattacatccatcataactga tgacatacat ctgacaggaa agattttctc atttttcaga 11760
 agtttccggcc accccagact tgaagcgtt aacggctgtc aaaaatgttag gaaatacatg 11820
 aatcagccata aagtcatgt gtatgagact ctgtatgaaag gtcatgccc attttgcgtt 11880
 atcataatca acggctatcg tgacaggac ggaggcgtt ggccacccgt gaccctcccc 11940
 ctgcgtctg cagacacaat ccggaaatgtt caagcttcag gtgaagggtt aacacatgag 12000
 cagtgcgtt ataaactggaa atctttgtt ggagtgttata ttggcgttgc tatgccttt 12060
 agcctggata gtgtatgttgc aatgtaccta aaggacaagg cacttgcgtc tctccaaagg 12120
 gaatgggatt cagtttaccc gaaagagttc ctgcgttacg accctcccaa gggAACGGG 12180
 tcacggggc ttgtatgttgc ttcccttaat gattcgatgttgc accctcccaa gggAACGGG 12240
 atgtatgttgc ttgtatgttgc ttcccttaat gattcgatgttgc accctcccaa gggAACGGG 12300
 aaagaaaagg agatcaagga aacaggtaga cttttgtata aatgtactt aaaaatgagg 12360
 gcatgccaag tgattgttgc aatgtatgttgc ttcccttaat gattcgatgttgc accctcccaa gggAACGGG 12420
 aatgggatgg ccaaggatgttgc gtcacgttgc actaaggac tccacactt agctgtctca 12480
 ggagtcgttgc aagatcttgc aagaaatgttgc agggggggc cagtcttaaa aacctactcc 12540
 cgaagcccgatccatcataactgttgc tccacacaat gtttgcgttgc tatgggttc 12600
 cctcaagtttgc ttggcgttgc ccaaggatgttgc gtcacgttgc actaaggac tccacactt agctgtctca 12660
 acagtcgttgc cttttatgttgc gactgtatgttgc aagaaatgttgc gagatgttgc 12720

-continued

accatcagct tggctgcaca gaggcttaat gagatttacg gattgccctc attttccag 12780
tggctgcata agaggcttga gacctctgtc ctgtatgtaa gtgaccctca ttgcgcgcgc 12840
gaccttgcac cccatatccc gttatataaa gtcggcaatc atcaaatctt cattaagtac 12900
cctatggag gtatagaagg gtattgttag aagctgtgga ccatcagcac cattccctat 12960
ctatacctgg ctgcttatga gagcggagta aggattgtt cgtagtgca aggggacaat 13020
cagacccatag ccgtaacaaa aagggtaccc agcacatggc cctacaaccc taagaaacgg 13080
gaagctgcta gagtaactag agattactt gtaatttta ggcaaaggct acatgatatt 13140
ggccatcacc tcaaggcaaa tgagacaatt gtttcatcac attttttgt ctattcaaaa 13200
ggaatataatt atgatgggct acttgtgtcc caatcactca agagcatcgc aagatgtgt 13260
ttctggtcag agactatagt tgatgaaaca agggcagcat gcagtaataat tgctacaaca 13320
atggctaaaa gcatcgagag aggttatgac cgtagcttg catattccct gaacgtccta 13380
aaagtgatac agcaaattct gatctcttt ggcttcacaa tcaattcaac catgacccgg 13440
gatgtatcata tacccttcct cacaacaac gacctttaa taaggatggc actgttgc 13500
gctcctatttggggatgaa ttatctgaat atgagcggc tgtttgcag aaacatcggt 13560
gatccagtaa catcatcaat tgctgatctc aagagaatga ttctcgccctc actaatgcct 13620
gaagagaccc tccatcaagt aatgacacaa caaccggggg actcttcatt cctagactgg 13680
gctagcgacc cttactcagc aaatcttgcata tggtccaga gcatcactag actcctcaag 13740
aacataactg caaggtttgt cctgatccat agtccaaacc caatgttaaa aggattattc 13800
catgatgaca gtaaagaaga ggacgagggc ctggcgccat tcctcatggc cagggatatt 13860
atagtaccta gggcagctca tgaaatcctg gatcatatgt tcacaggggc aagagagtct 13920
atggcaggca tgctggatac cacaaggc ttgattcgag ccagcatgag gaaggggggg 13980
ttaacctctc gagtgataac cagattgtcc aattatgact atgaacaatt cagagcagg 14040
atggtgctat tgacaggaag aaagagaaat gtcttcattt acaaagatgc atgttcagtg 14100
cagctggcgca gagctctaag aagccatatg tggcgaggc tagctcgagg acggctatt 14160
tacggccttg aggtccctga tgtaactagaa tctatgegag gccacccatt tcggogtcat 14220
gagacatgtg tcatctgcga gtgtggatca gtcaactacg gatggtttt tgccctcg 14280
ggttgcacac tggatgatac tgacaaggaa acatcattct tgagagttcc atatatttgt 14340
tctaccactg atgagagaac agacatgaag ctggcttcg taagagcccc aagtgcattc 14400
ttgcgatctg ctgttagaat agcaacagtg tactcatggg cttacggtga tgatgatgc 14460
tcttggaaac aagcctgggtt gttggcttagg caaaggccca atgtgagccct ggaggagct 14520
agggtgatca ctcccatctc aacttcgact aatttagcgc ataggttgcgatgc 14580
actcaagtga aatactcagg tacatccctt gtccgagttt cgaggtatac cacaatctcc 14640
aacgacaatc tctcatttgtt catatcagat aagaagggtt atactaactt tatataccaa 14700
caaggaatgc tccttagggtt ggggtttta gaaacattgt ttcgactcga gaaagatacc 14760
ggatcatctc acacggattt acatcttcac gtgcacccacg attgttgcgt gatcccgatg 14820
atagatcatc ccaggatacc cagctccgc aagcttagacg tgagggcaga gctatgtacc 14880
aacccattga tatatgataa tgcaccttta attgacagag atgcaacaag gctatacacc 14940
cagagccata ggaggcacct tggaaatattt gttacatggt ccacacccca actatatcac 15000
attttagctc agtccacacg actatctatg attgacactgg taacaaaatt tgagaaggac 15060
catatgaatg aaatttcagc tctcataggg gatgcacata tcaatagttt cataactgag 15120

-continued

ttctgtca tagagccaag attattcaact atctacttgg gccagtgtgc ggccatcaat 15180
 tggcatttg atgtacatta tcatalogcca tcagggaaat atcagatggg tgagctttg 15240
 tcatcggtcc tttctagaat gagcaaagga gtgttaagg tgcttgc当地 tgctctaagc 15300
 cacccaaaga tctacaagaa attctggcat tgtggattata tagagectat ccatggct 15360
 tcacttgatg ctcaaaactt gcacacaact gtgtcaaca tggtttacac atgctatatg 15420
 acctacccg acctgttgtt gaatgaagag ttagaagagt tcacatttctt cttgtgtgaa 15480
 agcgacgagg atgttagtacc ggacagattc gacaacatcc aggcaaaaca cttatgtgtt 15540
 ctggcagatt tttactgtca accagggacc tgccc当地 acca 15600
 gagaaatgtg cagttctaac cgaccatatac aaggcagagg ctatgttatac tccagcagga 15660
 ttttcgtgga acataaatcc aattattgtta gaccattact catgctctt gacttatctc 15720
 cggcgaggat ogatcaaaca gataagattt agagttgatc caggattcat tttcgacgcc 15780
 ctgcgtgagg taaatgtcag tcagggaaag atcggcagca acaacatctc aaatatgagc 15840
 atcaaggcatt tcagaccccc acacgatgtat gttgcaaaat tgctcaaaga tatcaacaca 15900
 agcaaggcaca atctcccat ttcagggggc aatctcgcca attatgaaat ccatgcttc 15960
 cgcagaatcg ggttgaactc atctgcttgc tacaagctt ttgagatatac aacattaatt 16020
 aggagatgcc ttgagccagg ggaggacggc ttgttcttgg gtgaggatgc gggttctatg 16080
 ttgatcaattt ataaggagat acttaaacta aacaagtgc tctataatag tggggttcc 16140
 gccaattcta gatctggta aagggatta gcaccctatc cctccgaatg tggcattgtc 16200
 gaacacagaa tgggagtagg taatattgtc aaagtgc当地 ttaacgggag gccc当地 16260
 acgtgggtag gcagtgtaga ttgctcaat ttcatagttt gtaatatccc tacctcttagt 16320
 gtggggttta tccattcaga tatagagacc ttgcctgaca aagataactat agagaagcta 16380
 gaggaaatgg cagccatctt atcgtggct ctgctcttgg gaaaaatagg atcaactatg 16440
 gtgattaagc ttatgcctt cagggggat ttgttgc当地 gatttataag ttatgttaggg 16500
 ttcattata gagaagtgaa ccttgtatac ccttagataca gcaacttcat atctactgaa 16560
 tcttattttgg ttatgacaga tctcaaggctt aaccggctaa tgaatcttga aaagattaag 16620
 cagcagataa ttgaatcatc tgtgaggact tcacctggac ttataggc当地 catcctatcc 16680
 attaagcaac taagctgcat acaagcaatt gtgggagacg cagtttagtag aggtgatatc 16740
 aatcctactc tgaaaaact tacacctata gaggcaggc当地 tgatcaatg cggggtggca 16800
 attaacggac ctaagctgtg caaaaattt atccaccatg atgttgc当地 agggcaagat 16860
 ggattgttta attctatact catcctctac agggagttgg caagattcaa agacaaccaa 16920
 agaagtcaac aaggatgtt ccacgcttac cccgtattgg taagtagcag gcaacgagaa 16980
 ctttatctta ggatcacccg caaatttgg gggcacattc ttcttactc cgggacacaga 17040
 aagttgataa ataagtttcc ccagaatctc aagtccggctt atctgataact agacttacac 17100
 cagaatatct tcgttaagaa tctatccaag tcagagaaac agattattat gacgggggt 17160
 ttgaaacgtg agtgggtttt taaggtaaca gtcaaggaga coaaagaatg gtataagtt 17220
 gtcggataca gtgcctgtat taaggactaa ttgttgc当地 tccggaaacc taatctgccc 17280
 ctaggtgggtt aggcattttt tgcaatataat taaagaaaac ttgaaaata cgaagttct 17340
 attcccaact ttgtctgggtg gccc当地 ctcgctggcg cccgctggc 17400
 aacattccga ggggaccgctc ccctcggtaa tggcgaatgg gacgcggccg atccggctgc 17460

-continued

taacaaagcc	cgaaaggaag	ctgagttggc	tgctgccacc	gctgagcaat	aactagcata	17520
accccttggg	gcctctaaac	gggtctttag	gggtttttt	ctgaaaggag	gaactatatc	17580
cggatcgccc	cgcaggtaacc	cagttttgt	tcccttagt	gagggttaat	ttcgagctt	17640
gctaatcat	ggtcatagct	gtttcctgt	tgaaattgtt	atccgtcac	aattccacac	17700
acatacggag	ccggaagcat	aaagtgtaaa	gcctggggtg	cctaattgagt	gagctaactc	17760
acattaattt	cgttgcgtc	actgcccgt	ttccagtcgg	gaaacctgtc	gtgccagctg	17820
cattaatgaa	tcggccaacg	cgcggggaga	ggcgggttgc	gtattggcg	ctttccgct	17880
tcctcgctca	ctgactcgct	gchgctcggtc	gttcggctgc	ggcgagcggt	atcagctcac	17940
tcaaaggccg	taatacggtt	atccacagaa	tcagggata	acgcaggaaa	gaacatgtga	18000
gcaaaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	cgttgtggc	gttttccat	18060
aggctccgccc	cccctgacga	gcatcacaaa	aatcgacgct	caagtcagag	gtggcgaaac	18120
ccgacaggac	tataaagata	ccaggcggtt	ccccctggaa	gtccctcggt	gchgctctcct	18180
gttccgacccc	tgccgcttac	cgatcacgt	tccgccttc	tcccttcggg	aagcgtggcg	18240
ctttctcata	gctcacgctg	taggtatctc	agttcggtgt	aggtcggtcg	ctccaagctg	18300
ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgctg	ccttatccgg	taactatcg	18360
cttgagtccttca	acccggtaag	acacgactta	tcgcccactgg	cagcagccac	tggtaacagg	18420
attagcagag	cgaggtatgt	aggcggtgt	acagagtct	tgaagtgggt	gcctaactac	18480
ggctacacta	gaaggacagt	atttggatc	tgcgcctgc	tgaagccagt	taccttcgga	18540
aaaagagttt	gtagctcttgc	atccggcaaa	caaaccacgg	ctggtagcgg	tggtttttt	18600
gtttcaagc	agcagattac	gchgagaaaa	aaaggatctc	aagaagatcc	tttgatctt	18660
tctacgggggt	ctgacgctca	gtggAACGAA	aactcacgtt	aagggttttt	ggtcatgaga	18720
ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	18780
taaagtatat	atgagtaaac	tttgtctgac	agttaccaat	gtttatcag	tgaggcacct	18840
atctcagcga	tctgtctatt	tcgttcatcc	atagttgcct	gactccccgt	cgtgtagata	18900
actacgatac	gggaggggctt	accatctggc	cccagtgtctg	caatgatacc	gchagaccca	18960
cgctcaccgg	ctccagattt	atcagcaata	aaccagccag	ccggaaaggcc	cgagcgcaga	19020
agtggccttg	caacttttac	cgcctccatc	cagtctattt	attgttgccg	ggaagctaga	19080
gtaagtagtt	cgccagttaa	tagttgcgc	aacgttgggt	ccattgtac	aggcatcg	19140
gtgtcacgt	cgtcgtttgg	tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	19200
gttacatgt	cccccatgtt	gtgcaaaaaa	gccccgttagct	ccttcgggtcc	tccgatcg	19260
gtcagaagta	agttggccgc	agtgttatca	ctcatggtta	tggcagcaact	gcataattct	19320
cttactgtca	tgccatccgt	aagatgtttt	tctgtgactg	gtgagttactc	aaccaagtca	19380
ttctgagaat	agtgttatgc	gcccacgt	tgtcttgc	ccggcgtcaat	acgggataat	19440
accgcgcac	atagcagaac	ttaaaaatgt	ctcatcattt	gaaaacgttc	ttcggggcga	19500
aaactctca	ggatcttacc	gctgttgaga	tccagttcg	tgttaacccac	tcgtgcaccc	19560
aactgatctt	cagcatctt	tactttcacc	agcggtttgt	ggtgagcaaa	aacaggaagg	19620
caaaatgccg	caaaaaagg	aataaggccg	acacggaaat	gttgaataact	cataactctc	19680
cttttcaat	attattgttgc	catatgttgc	ggttattgtc	tcatgagccg	atacatat	19740
gaatgttattt	agaaaaataa	acaaatagg	gttccgcgc	cattcccc	aaaagtgc	19798

-continued

<210> SEQ_ID NO 3
 <211> LENGTH: 5253
 <212> TYPE: DNA
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: This is artificially synthesized MSP1 d-190 3D7 sequence ORF

 <400> SEQUENCE: 3

atgaccgtcg	cgccggccgag	cgtgcccgcg	gegctgeccc	tcctcgggga	gctgecccgg	60
ctgctgtgc	tggtgtgtt	gtgcctgcgg	gcccgtgtgg	gatccgtgac	ccacgaatcc	120
tatcaggagc	tggtaagaa	actgaaagct	ttagaggacg	ccgtattgac	aggttactcc	180
ctattccaga	aagaaaagat	ggttttaaac	gaagaagaaa	ttaccacaaa	gggagcatcc	240
gcccagtctg	gagcatctgc	tcagagcgg	gcatctgctc	agagtgggac	aagcgcccaa	300
agtggagcgt	ctgcccagtc	aggcgccctca	gctcaatctg	gaacctctgg	gccgagtggt	360
cctagcggta	cttctccaag	tagccgtct	aatacactcc	cacgttccaa	cacctccagt	420
ggagcctccc	caccggccga	cgcacccgac	tcagacgcta	agagttatgc	agacctgaag	480
caccgcgtga	ggaactacct	tttcaactatc	aaagagttga	agtaccctga	attgttcgat	540
ttgaccaacc	atatgctgac	actctgtgac	aacatacatg	gtttcaagta	tctgatagat	600
gggttatgaag	aaattaacga	gctgctctat	aaactcaact	tttacttcga	cctgctgcgt	660
gccaagctga	acgatgtctg	tgcaaacgat	tactgccaga	tcccattcaa	cctaaagata	720
cgtgcgaacg	agctggatgt	tctgaagaaa	ctcgtgttcg	ggtatcgaa	acccttgac	780
aacattaagg	acaatgtggg	gaagatggag	gattacatta	agaaaaataa	aacaacaatc	840
gctaacataa	atgagcttat	cgaggggac	aaaaagacca	tcgaccagaa	caagaatgcc	900
gacaatgaag	agggaaaaaa	gaaactatac	caagccccagt	atgatttgag	catctacaat	960
aagcaactag	aggaagctca	caacccatc	agcgtactgg	aaaagagaat	tgacaccctg	1020
aaaaagaatg	aaaacattaa	gaaactccctg	gacaagatta	acgaaatcaa	aaacccacct	1080
ccagcgaata	gcggaaatac	cccgaaatacc	ctgctggata	agaacaaaaa	gattgaagag	1140
cacgaagaga	aaatcaagga	aatcgccaag	actattaagt	tcaatataga	ttctctgttc	1200
acagaccctc	tggagctgga	atactacctg	cgcgagaaga	ataagaaggt	cgacgtgacc	1260
ccaaagagcc	aagacccaaac	aaagtccctg	cagatcccc	aagtgcctca	cccaaacggc	1320
atcgtgtatc	ccctgcctct	taccgacatc	cacaactctc	tggcagccga	taacgacaaa	1380
aacagctatg	gagacctgat	gaaccccccac	actaaggaaa	agataaaacga	gaagatcatt	1440
accgataata	aggagcggaa	gatttttac	aacaacatca	agaagaaaat	cgacctggaa	1500
gagaaaaata	tcaatcacac	caaagagcaa	aacaagaaaat	tactggagga	ctatgagaag	1560
agcaaaaagg	attatgagga	actgttagag	aagttctatg	aatgaaatt	caacaacaat	1620
ttcgataagg	atgtggtcga	taaaaatttc	agcgccccgt	acacctacaa	cgtggagaag	1680
cagcggtaca	acaataagt	cagcagctcc	aataactcg	tctacaatgt	gcagaagctg	1740
aagaaagctc	ttagctatct	ggaagactac	tcgctgagga	aagggatttc	tgagaaggat	1800
ttcaaccact	actacaccct	caaaacccgc	ctggaagctg	acatcaagaa	actcactgaa	1860
gagatcaaaa	gttctgagaa	taagatactg	gagaagaact	tcaagggact	aacgcactct	1920
gaaaacggct	ccctggaaat	ctctgacatc	gtgaaactgc	aagtccaaa	ggtgtctgc	1980
atcaaaaaaa	tcgaggatct	gcfgaaagatc	gagctgtttc	ttaagaacgc	ccaaactgaaa	2040
gactcaatcc	acgtgcctaa	catttacaaa	ccgcagaaca	aaccagaacc	atactatctg	2100

-continued

atcgtgctga agaaggaggt ggataagctg aaggaattca tccaaaagt gaaagatatg	2160
ttaaagaag agcaagccgt gctgagcgc ataacgcgc ctctggggc cgcaagcgag	2220
acaaccgaag atggggggca cagcacccac accctgtctc agtctggcga aacagaggtg	2280
acagaagaga cagaagagac cgaagaaaca gtggggcaca ccactactgt gaccatcact	2340
ttggcccccta cgccggccatc tccccaaaa gaggtcaaag tcgtggaaaa ctccattgaa	2400
cagaagtcca acgacaactc acaggctcg acgaagaccg tctatctgaa gaaactggac	2460
gagttctgaa caaaaagcta catctgccc aaatacatcc tcgtgtctaa cagcagcatg	2520
gatcagaagc tggtggaggt gtacaaccta acgccccgaag aagagaacga gttaaaatcc	2580
tgtgatccct tagacctact gtttaacatt cagaacaaca tccccgctat gtacagctt	2640
tatgattcca tgaataacga cctccagcac ctgttctcg agctgtacca gaaagagatg	2700
atctactatc tgcataagct gaaagaggag aatcacatca aaaagttgct ggaagagcag	2760
aaacagataa ctggggcgtc cagcacatcg tcacctggca acacgacagt aaataccgcc	2820
cagtcgtcta cacactccaa ctcggcggac cagcagagca acgcttctag caccacacc	2880
cagaatgggg tagcagttag tagggccct gctgtgggtt aggaatcgca tgacccttc	2940
actgtattat otatttcaaa cgacctaaaa gggattgtgt coctctcaa tttaggtaat	3000
aagaccaagg tccctaaccctt gactatc acgactacgg aaatggagaa gttttatgaa	3060
aacatectgaa agaacaacga cacctatccc aacgacgaca taaagcagtt cgtgaagagt	3120
aacagtaaag tgattaccgg gctgacagaa acccagaaaa atgctttaaa tgatgagatc	3180
aagaaaactgaa aagacacact ccagctctcc ttcgatctgt acaacaagta caaactaaag	3240
ctggacagat tattcaataa gaagaaggag ctggggcaag ataagatgca gattaagaag	3300
ctaactttac tgaaggagca gctcgagac aagctcaact ccctgaataa tccacataat	3360
gtgctccaga actttccgtt attcttcaat aagaagaaag aagcagagat tgccgagacg	3420
aaaaataccctcgaa tcgaaaacac taagatatta ctgaaaacact ataaaggct ggtgaagttat	3480
tacaacggag agtcttagccc attgaagact ctttcagaag tgtcaattca aaccgaggat	3540
aactacgcaaa acctagaaaaa gttcagagtg ctgagcaaaa tcgacggcaa actcaatgt	3600
aacctacacc tcggaaaaaa aaagctgac ttctgtccca gtggacttca tcatttaatt	3660
accgaatttga aagaagttat caaaaacaaa aactacactg ggaacagccc atctgaaaat	3720
aataaaaagg tcaacggggc cctcaagtct tatgaaaattt tccttccaga agcaaaagt	3780
acaaccgtcg tgacccccc ccagcccgat gtcaccccca gccctctaag cgtgagagtg	3840
tctggatcaa gtgggtccac aaaagaagaa acccagatcc ccacatcagg atctctactg	3900
accgagttgc agcaggctcg ccaactcccg aattatcgac aggaagacgca cagccctcg	3960
gttttgcataa tcttcggcga atcagaagac aacgacgagt accttagacca agtgggtcacc	4020
ggggaaagcga ttatgttcac tatggacaatt atcctcagcg gtttcgagaa cgagtatgac	4080
gtgatctacc tcaaaccact agccggagtt tacagaagtc tcaagaagca gatcgaaaag	4140
aacatcttca cttttaatctt aaacccaaac gacatcttgc attcccgctt gaaaaagcgg	4200
aaatacttcc tcgactact ggagtcggat ttgtatgcgt ttaagcacat ctccagcaac	4260
gaatacatttc tcgaggactc gttcaactg ttaaactccg agcagaagaa caccctgtcg	4320
aagtccataa aatataatcaa agagtcagtc gagaacgata ttaaattcgc ccaagaaggc	4380
ataagctact acgaaaaggt cctcgccaaa tacaaggacg atctggagtc tatcaaaaag	4440

-continued

gtcatcaaag aagagaaaaga gaaattcccc agttctcccc ctacaacgcc gcccctccca	4500
gccaagactg atgaacagaaa aaaagagtct aagttccctc ctttcctcac taatatcgag	4560
actctctaca ataaccttagt gaacaagatt gacgactacc tgatcaacct taaagccaa	4620
ataaacgact gcaatgtcga gaaggatgag gctcatgtta agatcaccaa actgtccgat	4680
ctgaaagcca tcgacgacaa gatcgactta tttaaaaacc catacgattt cgaggctatc	4740
aaaaagctga tcaatgtatca caccaagaaa gatatgtcg gcaagctgt gaggcacgggt	4800
ctgggtcaga acttccctaa caccatcata tcaaagctca tagagggcaa gttccaagac	4860
atgctgaata tttcacagca tcagtgcgtc aagaagcagt gccccaaaa ttctggatgc	4920
ttccggcacc tggatgagcg agaagagtgc aagtgcgtgc ttaactataa acaggaggc	4980
gacaaatgtg tggagaaccc aaatcccgacg tgcaacgaga acaacgggtt ctgcgtatgc	5040
gacgcgactt gtacagagga agactcggtt agttctcgaa aaaaaatcac gtgcgtatgc	5100
accaaaccgg acagttatcc tctgttcgtat gggatattct gtcctccag caacgttact	5160
accccccggca ctaccggcgtct tctatctggt cacacgtgtt tcacgttgcg aggttgtt	5220
gggacgctag taaccatggg cttgtact taa	5253

<210> SEQ_ID NO 4

<211> LENGTH: 5160

<212> TYPE: DNA

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: This is artificially synthesized MSP1 d-190*
3D7sequence ORF

<400> SEQUENCE: 4

atgaccgtcg cgccggccgag cgtcccccgcg gcgctgcccc tccctggggg gctccccgg	60
ctgctgtgc tggtgctgtt gtgcgtgcgg gccgtgtggg gatccgtgac ccacgaatcc	120
tatcaggagc tggtaagaa actggaaatc ttagaggacg ccgtatttgcg aggttactcc	180
ctattccaga aagaaaagat ggtttaaac gaagaagaaa ttaccacaaa gggagcatcc	240
gcccagtctg gacgtctgc tcagagcgga gcatctgtctc agatggggc aagcgccaa	300
agtggagcgt ctgcccagtc agggccctca gctcaatctg gaacctctgg gccgagtgg	360
cctageggta cttctccaag tagegggtct aatacactcc cacgttccaa cacctccagt	420
ggagccccc caccggccga cgcatccgac tcagacgtta agatgtatgc agacctgaag	480
caccgggtga ggaactaccc tttcaacttc aaagagtgtt agtaccctgtt attgttcgtat	540
ttgaccaacc atatgtgtac actctgtgac aacatacatg gtttcaagta tctgtatagat	600
gggttatgtt aatattaacga gctgtcttat aaactcaact tttacttcga cctgtgtcg	660
gccaagctga acgtatgtctg tgcaaaccgt tactgccaga tccccattcaa cctaaagata	720
cgtgcgaacg agctggatgt tctgtatcc ggtatcgaa acccttggac	780
aacattaagg acaatgtggg gaagatggag gattacatta agaaaaataa aacaacaatc	840
gctaacataa atgagcttac cgagggggc aaaaagacca tcgaccagaa caagaatgcc	900
gacaatgaag agggaaaaaaa gaaactatac caagccagt atgatttgag catctacaat	960
aagcaactag aggaagctca caacctcatc agcgtaactgg aaaagagaat tgacaccctg	1020
aaaaagaatg aaaacattaa gaaactccgt gacaagatca acgaaatcaa aaacccacct	1080
ccagcgaata gcggaaatac cccgaatacc ctgctggata agaacaaaaa gattgaagag	1140
cacgaagaga aatcaagga aatcgccaag actattaagt tcaatataga ttctctgttc	1200

-continued

acagaccctc tggagctgga atactacctg cgcgagaaga ataagaagg cgacgtgacc	1260
ccaaagagcc aagacccaac aaagtccgtg cagatcccc aagtgccta cccaaacggc	1320
atcgtgtatc ccctgctct tacggacatc cacaactctc tggcagccga taacgacaaa	1380
aacagctatg gagacctgat gaacccccc actaaggaaa agataaacga gaagatcatt	1440
accgataata aggagccgaa gattttatac aacaacatca agaagaaaat cgacctgaa	1500
gagaaaaata tcaatcacac caaagagcaa aacaagaaa tactggagga ctatgagaag	1560
agcaaaaagg attatgagga actgttagag aagttctatg aaatgaaatt caacaacaat	1620
ttcgataagg atgtggcga taaaatttc agcgcccggt acacctacaa cgtggagaag	1680
cagcgttaca acaataagtt cagcagctcc aataactcggt tctacaatgt gcagaagctg	1740
aagaaagctc tgagctatct ggaagactac tcgctgagga aagggattc tgagaaggat	1800
ttcaaccact actacaccct caaaacccgc ctggaaagctg acatcaagaa actcactgaa	1860
gagatcaaaa gttctgagaa taagatactg gagaagaact tcaagggact aacgcactct	1920
gcaaaccggct ccctggaagt ctctgacatc gtgaaactgc aagtccaaa ggtgctgctc	1980
atcaaaaaaa tcgaggatct gcgaaagatc gagctgtttc ttaagaacgc ccaactgaaa	2040
gactcaatcc acgtgcctaa catttacaaa ccgcagaaca aaccagaacc atactatctg	2100
atcgtgctga agaaggaggt ggataagctg aaggaattca tcccaaaagt gaaagatatg	2160
ttaaagaaag agcaagccgt gctgagcgc ataacgcgc ctctggtggc cgcaagcgc	2220
acaaccgaag atggggggca cagcacccac accctgtctc agtctggcga aacagaggt	2280
acagaagaga cagaagagac cgaagaaaca gtggggcaca ccactactgt gaccatcact	2340
ttgcccccta cgcagccatc tcccccaaaa gaggtcaaag tcgtggaaaa ctccattgaa	2400
cagaagtcca acgacaactc acaggctcg acgaagaccg tctatctgaa gaaactggac	2460
gagttcctga ccaaagcta catctgccat aaatacatcc tcgtgtctaa cagcagcatg	2520
gatcagaagc tggggggatgttacaccta acgcccgaag aagagaacga gttaaatcc	2580
tgtgatccct tagacctact gttiacatt cagaacaaca tccccgtat gtacagctt	2640
tatgattcca tgaataacga cctccagcac ctgttctcg agctgtacca gaaagagatg	2700
atctactatc tgcataagct gaaagaggag aatcacatca aaaagttgt ggaagagcag	2760
aaacagataa ctgggcgtc cagcacatcg tcacctggca acacgcacgt aaataccgccc	2820
cagtctgcta cacactccaa ctcccagaac cagcagagca acgcttctag caccacacc	2880
cagaatgggg tagcagttag tagcgccct gctgtgggg aggaatcgca tgaccctc	2940
actgtattat ctattcaaa cgacctaaaa gggattgtgt ccctcctcaa ttttaggtat	3000
aagaccaagg tccctaaccctt cttgactatc agcactacgg aaatggagaa gtttatgaa	3060
aacatccctga agaacaacga cacctatccc aacgacgaca taaagcagg tctgtggat	3120
aacagtaaag tgattaccgg gctgacagaa acccagaaaa atgcattaa tgatgagatc	3180
aagaaaactga aagacacact ccagctctcc ttgcgtatgt acaacaagta caaactaaag	3240
ctggacagat tattcaataa gaagaaggag cttggcgaag ataagatgca gattaagaag	3300
ctaactttac tgaaggagca gctcgagacg aagctcaact ccctgaataa tccacataat	3360
gtgctccaga actttccgt attcttcaat aagaagaaag aagcagagat tgccgagacg	3420
gaaaataccc tcgaaaacac taagatatta ctgaaacact ataaagggt ggtgaagat	3480
tacaacggag agtcttagccc attgaagact ctttcagaag tgtcaattca aaccgaggat	3540
aactacgcaaa acctagaaaa gttcagagtg ctgagcaaaa tcgacggcaa actcaatgtat	3600

-continued

aacctacacc tcggaaaaaa aaagctgagc ttccgttcca gtggacttca tcatattaatt 3660
 accgaattga aagaagttat caaaaacaaa aactacactg ggaacagccc atctgaaaat 3720
 aataaaaagg tcaacgaggc cctcaagtct tatgaaaatt tcttccaga agcaaaagt 3780
 acaaccgtcg tgacccccc ccagccccat gtcaccccca gccctctaag cgtgagatg 3840
 tctggatcaa gtggctccac aaaagaagaa acccagatcc ccacatcagg atctctactg 3900
 accgagttgc agcaggtcgt ccaactccag aattatgacg aggaagacga cagcctcgt 3960
 gtttgccaa tcttggcga atcagaagac aacgacgagt accttagacca agtggtcacc 4020
 ggggaagcga ttagtgtcac tatggacaat atcctcagcg gttcgagaa cgagtatgac 4080
 gtgatctacc tcaaaccact agccggagtt tacagaagtc tcaagaagca gatcgaaaag 4140
 aacatcttca cttaaatct aaacctaaac gacatcttga attcccggt gaaaaagcgg 4200
 aaatacttcc tcgacgtact ggagtcggat ttgatgcagt ttaagcacat ctccagcaac 4260
 gaatacatta tcgaggactc gttcaactg ttaaactccg agcagaagaa caccctgctg 4320
 aagtccata aatatataa agagttagtc gagaacgata ttaaattcgc ccaagaaggc 4380
 ataagctact acgaaaaggt cctcgccaa tacaaggacg atctggagtc tatcaaaaag 4440
 gtcataaag aagagaaaga gaaatttccc agttctccc ctacaacgcc gccctctcca 4500
 gccaagactg atgaacagaa aaaagagtct aagttccctc ctttcctcac taatatcgag 4560
 actctctaca ataacctagt gaacaagatt gacgactacc tcatcaacct taaagccaaag 4620
 ataaacgact gcaatgtcgaa gaaggatgag gctcatgttta agatcaccaa actgtccgat 4680
 ctgaaagcca tcgacgacaa gategactta tttaaaacc catacgatt cgaggctatc 4740
 aaaaagctga tcaatgtga caccaagaaa gatatgtcg gcaagctgct gaggcacgggt 4800
 ctggcaga acttcctaa caccatcata tcaaagctca tagagggcaa gttccaagac 4860
 atgctgaata tttcacagca tcagtgcgtc aagaaggcgt gccccaaaa ttctggatgc 4920
 ttccggcacc tggatgagcg agaagagtgc aagtgcgtc ttaactataa acaggaggc 4980
 gacaaatgtg tggagaaccc aaatccgacg tgcaacgaga acaacgggtt ctgcgtatgc 5040
 gacgcgactt gtacagagga agactcgggg agttctcgaa aaaaatcac gtgcgtatgc 5100
 accaaacccg acagttatcc tctgttcgtat gggatattct gtcctccag caacgtttag 5160

<210> SEQ ID NO 5
 <211> LENGTH: 4122
 <212> TYPE: DNA
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: This is artificially synthesized MSP1 d-83-30-38 3D7 sequence ORF

<400> SEQUENCE: 5

atgaccgtcg cgccggccgag cgtccccggc gegctgcccc tctcgaaaaa gctgccccgg 60
 ctgctgtgc tggatgtgtt gtgcgtgcgg gccgtgtgg gatccgtgac ccacgaatcc 120
 tatcaggagc tggatgttggaa actggaaatc tttagggacg ccgtatttgc aggttactcc 180
 ctatccaga aaaaaatgttggaaatc gatggatggaaatc ttaccacaaaa gggagcatcc 240
 gcccagtctg gagcatctgc tcagagcgaa gcatctgcgtc agagtggagc aagcgccaa 300
 agtggagcgt ctgcggccatc aggccgttca gctcaatctg gaaaccttgg gccgtgtgtt 360
 cttctccaaatc tagccgttca aatacactcc cacgttccaa cacctccatgtt 420
 ggagccccc caccggccgaa cgcatccgac tcagacgctaa agagttatgc agacgttgc 480

-continued

caccgcgtga ggaactacct tttcactatac aaagagttga agtaccctga attgttcgat	540
ttgaccaacc atatgtcgac actctgtgac aacatacatg gtttcaagtg tctgatagat	600
gggttatgaag aaattaacga gctgctctat aaactcaact ttacttcga cctgctgcgt	660
gccaagctga acgatgtctg tgcaaacgat tactgccaga tcccattcaa cctaaagata	720
cgtgcgaacg agctggatgt tctgaagaaa ctcgttccg ggtatcgaa acccttggac	780
aacattaagg acaatgtggg gaagatggag gattacatta agaaaaataa aacaacaatc	840
gctaacataa atgagcttat cgaggggagc aaaaagacca tcgaccagaa caagaatgcc	900
gacaatgaag agggaaaaaaa gaaactatac caagcccagt atgatttgag catctacaat	960
aagcaactag aggaagctca caacctcatac agcgtactgg aaaagagaat tgacaccctg	1020
aaaaagaatg aaaacattaa gaaactcctg gacaagatta acgaaattaa aaacccacct	1080
ccagcgaata gcggaaatac cccgaatacc ctgctggata agaacaaaaa gattgaagag	1140
cacgaagaga aaatcaagga aatcgccaag actattaagt tcaatataga ttctctgttc	1200
acagacoctc tggagctgga atactacctg cgcgagaaga ataagaaggt cgacgtgacc	1260
ccaaagagcc aagacccaac aaagtccctg cagatcccc aagtgcocca cccaaacggc	1320
atcgtgtatc ccctgcctct tacccacatc cacaactctc tggcagocga taacgacaaa	1380
aacagctatg gagacctgat gaaccccccac actaaggaaa agataaacga gaagatcatt	1440
accgataata aggagcggaa gatTTTATC aacaacatca agaagaaaat cgacctggaa	1500
gagaaaaata tcaatcacac caaagagca aacaagaaaat tactggagga ctatgagaag	1560
agcaaaaagg attatgagga actgttagag aagttctatg aaatgaaatt caacaacaat	1620
ttcgataagg atgtggcga taaaatttc agcgcgggat acacctacaa cgtggagaag	1680
cagcgttaca acaataagtt cagcagctcc aataactcgg tctacaatgt gcagaagctg	1740
aagaaagctc ttagctatctt ggaagactac tcgctgagga aagggatttc tgagaaggat	1800
ttaaaccact actacacccct caaaacccgc ctggaagctg acatcaagaa actcactgaa	1860
gagatcaaaa gttctgagaa taagatactg gagaagaact tcaaggact aacgactct	1920
gcaaaacggct ccctggaaatg ctctgacatc gtgaaactgc aagtccaaaa ggtgctgctc	1980
atcaaaaaaa tcgaggatct gcgaaagatc gagctgtttc ttaagaacgc ccaactgaaa	2040
gactcaatcc acgtgcctaa catttacaaa ccgcagaaca aaccagaacc atactatctg	2100
atcgtgtga agaaggaggt ggataagctg aaggaattca tcccaaaatg gaaagatatg	2160
ttaaagaaag agcaagccgt gctgagcagc ataacgcagc ctctggcgc cgcaagcag	2220
acaacccgaag atggcgggca cagcacccac accctgtctc agtctggcga aacagaggt	2280
acagaagaga cagaagagac cgaagaaaca gtggggcaca ccactactgt gaccatcact	2340
ttggcccccata cgccagccat tcccccaaaa gaggtcaaaatg tcgtggaaaa ctccattgaa	2400
cagaagtcca acgacaactc acaggctctg acgaagacgg tctatctgaa gaaactggac	2460
gagttcctga ccaaaagcta catctgccat aaatacatcc tcgtgtctaa cagcagcatg	2520
gatcagaagc tggggaggt gtacaaccta acgcccgaag aagagaacga gttaaaatcc	2580
tgtgatccct tagacctact gtttaacatt cagaacaaca tccccgctat gtacagctt	2640
tatgattcca tgaataacga cctccagcac ctgttctcg agctgtacca gaaagagatg	2700
atctactatc tgcataagct gaaagaggag aatcacaatca aaaagttgct ggaagagcag	2760
aaacagataa ctgggacgatc cagcacatcg tcacctggca acacgacagt aaataccgcc	2820

-continued

cagtctgcta cacactccaa ctcccagaac cagcagagca acgcttctag caccacacc	2880
cagaatgggg tagcagttag tageggccct gctgtgggtgg aggaatcgca tgacccttc	2940
actgtattat ctatttcaa cgacctaaaa gggattgtgt ccctctcaa ttttagtaat	3000
aagaccaagg tccctaaccctt cttgactatc agcactacgg aaatggagaa gtttatgaa	3060
aacatcctga agaacaacga cacctatTTT aacgacgaca taaagcaggT cgtgaagagt	3120
aacagtaaag tgattaccgg gctgacagaa acccagaaaa atgcTTaaa tgatgagatc	3180
aagaaaactga aagacacact ccagctctcc ttgcatactgt acaacaagta caaactaaag	3240
ctggacagat tattcaataa gaagaaggag cttggcaag ataagatgca gattaagaag	3300
ctaactttac tgaaggagca gctcgagac aagctcaact ccctgaataa tccacataat	3360
gtgctccaga acttttcgtt attcttcaat aagaagaaAG aagcagagat tgccgagacg	3420
gaaaatacccc tcgaaaacac taagatatta ctgaaacact ataaagggtt ggtgaagtt	3480
tacaacggag agtctagccc attgaagact ctttcagaag tgtcaattca aaccgaggat	3540
aactacgcaa acctagaaaa gttcagagtg ctgagcaaa tcgacggccaa actcaatgt	3600
aacctacacc tcggaaaaaa aaagctgaccc ttccctgtcca gtggacttca tcatttaatt	3660
accgaattga aagaagttat caaaaacaaa aactacactg ggaacagccc atctgaaaat	3720
aataaaaagg tcaacgaggc cctcaagtct tatgaaaatt tccttccaga agcaaaagt	3780
acaaccgtcg tgacccccc ccagcccgat gtcaccccca gcccctctaag cgtgagatg	3840
tctggatcaa gtggctccac aaaagaagaa acccagatcc ccacatcagg atctctactg	3900
accgagttgc agcaggtcgt ccaactccag aattatgacg aggaagacga cagcctcg	3960
gttttgcCAA tcttcggcGA atcagaagac aacgacgagt accttagacca agtggtcacc	4020
aacgttacta ttccggcac tacccgttt ctatctggc acacgtgtt cacgttgaca	4080
ggtttgcttg ggacgctagt aaccatgggc ttgctgactt aa	4122

<210> SEQ_ID NO 6
 <211> LENGTH: 4029
 <212> TYPE: DNA
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: This is artificially synthesized MSP1 d-83-30-
 38* 3D7 sequence ORF

<400> SEQUENCE: 6	
atgaccgtcg cgccggccgag cgtgcccgcg ggcgtgcccc tcctcgggg gctgccccgg	60
ctgctgtgc tggcgtgtt gtgcctgccc gccgtgtggg gatccgtgac ccacgaatcc	120
tatcaggagc tggtaagaa actggaagct ttagaggacg ccgtattgac agttactcc	180
ctatccaga aagaaaagat ggtttaaac gaagaagaaa ttaccacaaa gggagcatcc	240
gcccagtctg gacgtctgc tcagagcggaa gcatctgctc agagtgggc aagcgccaa	300
agtggagcgt ctgcccagtc aggccctca gctcaatctg gaacctctgg gccgagtgg	360
cctagcggtt cttctccaa tagccggct aatacactcc cacgttccaa cacctccagt	420
ggagcctccc cacccggcga cgcacccgac tcagacgcta agagttatgc agacgtgaa	480
caccgcgtga ggaactaccc ttctactatc aaagagtgtt agtaccctga attgttcgt	540
ttgaccaacc atatgtgtac actctgtgac aacatacatgt gttcaagta tctgtatgt	600
gggttatgtaa aatTAACGA gctgctctat aaactcaact ttacttcga cctgtcg	660
gccaagctga acgtatgtctg tgcaaacatg tactgccaga tccattcaa cctaaagata	720

-continued

cgtgcgaacg agctggatgt tctgaagaaa ctcgtgttcg ggtatcgaa acccttggac	780
aacattaagg acaatgtggg gaagatggag gattacatta agaaaaataa aacaacaatc	840
gctaacataa atgagottat cgaggggagc aaaaagacca tcgaccagaa caagaatgcc	900
gacaatgaag agggaaaaaa gaaactatac caagcccagt atgatttgag catctacaat	960
aagcaactag aggaagctca caacctcatc agcgtaactgg aaaagagaat tgacaccctg	1020
aaaaagaatg aaaacattaa gaaactccctg gacaagatta acgaaattaa aaacccacct	1080
ccagcgaata gcggaaatac cccgaatacc ctgctggata agaacaaaaa gattgaagag	1140
cacgaagaga aaatcaagga aatcgccaag actattaagt tcaatataga ttctctgttc	1200
acagaccctc tggagctgga atactacctg cgcgagaaga ataagaaggt cgacgtgacc	1260
ccaaagagcc aagacccaac aaagtccctg cagatcccc aagtccctc cccaaacggc	1320
atcgtgtatc ccctgcctct taccgacatc cacaactctc tggcagccga taacgacaaa	1380
aacagctatg gagacctgat gaacccccc actaaggaaa agataaacga gaagatcatt	1440
acccgataata aggagccgaa gatttttatc aacaacatca agaagaaaat cgacctggaa	1500
gagaaaaata tcaatcacac caaagagcaa aacaagaaat tactggagga ctatgagaag	1560
agcaaaaagg attatgagga actgttagag aagttctatg aaatgaaatt caacaacaat	1620
ttcgataagg atgtggtcga taaaattttc agcgccccgt acacctacaa cgtggagaag	1680
cagcggtaca acaataagtt cagcagctcc aataactcg tctacaatgt gcagaagctg	1740
aagaaagctc tgagctatct ggaagactac tcgctgagga aagggatttc tgagaaggat	1800
ttcaaccact actacacccct caaaaccggc ctggaagctg acatcaagaa actcactgaa	1860
gagatcaaaa gttctgagaa taagatactg gagaagaact tcaagggact aacgcactct	1920
gcaaacggct ccctggaaatg ctctgacatc gtgaaactgc aagtccaaa ggtgctgctc	1980
atcaaaaaaa tcgaggatct gcgaaagatc gagctgtttc ttaagaacgc ccaactgaaa	2040
gactcaatcc acgtgcctaa catttacaaa ccgcagaaca aaccagaacc atactatctg	2100
atcgtgctga agaaggaggt ggataagctg aaggaattca tcccaaaagt gaaagatatg	2160
ttaaagaaag agcaagccgt gctgagcgc ataacgcgc ctctgggatc cgcaagcgag	2220
acaacccgaag atggccggca cagcacccac accctgtctc agtctggcga aacagaggtg	2280
acagaagaga cagaagagac cgaagaaaca gtggggcaca ccactactgt gaccatcaet	2340
ttgcccccta cgcagccatc tcccccaaaa gaggtcaaag tcgtggaaaa ctccattgaa	2400
cagaagtcca acgacaactc acaggctcg acgaagaccg tctatctgaa gaaactggac	2460
gagttcctga ccaaaagctt catctgccccat aaatacatcc tcgtgtctaa cagcagcatg	2520
gatcagaagc tggggggatgt gtacaaccta acgccccggaa aagagaacga gttaaaatcc	2580
tgtgatccct tagacctact gtttaacatt cagaacaaca tccccgttat gtacagctt	2640
tatgattcca tgaataacga cctccagcac ctgttcttcg agctgtacca gaaagagatg	2700
atctactatc tgcataagct gaaagaggag aatcacatca aaaagtgtt ggaagagcag	2760
aaacagataa ctgggacgtc cagcacatcg tcacctggc acacgacagt aaataccggcc	2820
cagtctgcta cacactccaa ctcccagaac cagcagagca acgcttcttag caccaacacc	2880
cagaatgggg tagcagttt tagcggccct gctgtggatgg aggaatcgca tgacccctc	2940
actgttattat ctatccaaa cgacccaaaa gggattgtgt ccctccctaa ttttaggtat	3000
aagaccaagg tcccttaaccc cttgactatc agcactacgg aatggagaa gttttatgaa	3060
aacatccctga agaacaacga cacctatccc aacgacgaca taaagcagtt cgtgaagagt	3120

-continued

aacagtaaaag tgattaccgg gctgacagaa acccagaaaa atgcttaaa tcatgagatc	3180
aagaaaactga aagacacact ccagctctcc ttgcgtatgt acaacaagta caaactaaag	3240
ctggacagat tattcaataa gaagaaggag ctggggcaag ataagatgca gattaagaag	3300
ctaactttac tgaaggagca gctcgagac aagctcaact ccctgaataa tccacataat	3360
gtgctccaga actttccgt attcttcaat aagaagaaag aagcagagat tgccgagacg	3420
gaaaataccc tcgaaaacac taagatatta ctgaaacact ataaagggtt ggtgaagtat	3480
tacaacggag agtcttagccc attgaagact ctttcagaag tgtcaattca aaccgaggat	3540
aactacgcaa acctagaaaa gttcagagtg ctgagcaaaa tcgacggcaa actcaatgat	3600
aacctacacc tcggaaaaaa aaagctgagc ttctgtcca gtggacttca tcatttaatt	3660
accgaattga aagaagttat caaaaacaaa aactacactg ggaacagccc atctgaaaat	3720
aataaaaagg tcaacgaggc cctcaagtct tatgaaaatt tccttcaga agcaaaagtg	3780
acaaccgtcg tgaccccccc ccagcccat gtcacccccc gcctcttaag cgtgagatg	3840
tctggatcaa gtggctccac aaaagaagaa acccagatcc ccacatcagg atctctactg	3900
accgagttgc agcaggttgtc ccaactccag aattatgacg aggaagacga cagcctctg	3960
gttttgocaa ttctggcga atcagaagac aacgacgagt accttagacca agtggtcacc	4020
ggggataaa	4029

<210> SEQ ID NO 7
 <211> LENGTH: 1347
 <212> TYPE: DNA
 <213> ORGANISM: Artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: This is artificially synthesized MSP1 d-42 3D7 sequence ORF

<400> SEQUENCE: 7

atgaccgtcg cgccggccgag cgtcccccg gcgctgcccc tcctcgaaaa gctccccgg	60
ctgctgtgc tgggtgtgtt gtgcctgccc gccgtgtggg gatccgtggt caccggggaa	120
gcgatttagtg tcactatggc caaatatcc tcgccccctcg agaacggatc tgacgtgatc	180
tacctcaaac cactagccgg agtttacaga agtctcaaga agcagatcga aaagaacatc	240
ttcaccttta atctaaaccc aaacgacatc ttgaattccc ggctgaaaaa gcggaaatac	300
ttcctcgacg tactggatgc ggatttgatc cagtttaagc acatctccag caacgaatac	360
attatcgagg actcggttcaa actgtttaac tccgagcaga agaacaccc gctgaagtcc	420
tacaaatata tcaaagatgc agtcgagaac gatattaaat tcgccccaga aggctaaac	480
tactacgaaa aggtccctcg caaatacaag gacgatctgg agtcttatcaa aaagggtcatc	540
aaagaagaga aagagaaatt tcccagttct cccccataaa cggccgcctc tccagccaa	600
actgtatgaa agaaaaaaa gtcatagttc ctcctttcc tcactaatat cgagacttc	660
tacaataacc tagtgaacaa gattgacgac tacctgtatca accttaaagc caagataaac	720
gactgcaatg tcgagaagga tgaggctcat gttaagatc ccaaaactgtc cgatctgaaa	780
gccatcgacg acaagatcga cttatataa aacccatacg atttcgaggg tatcaaaaag	840
ctgatcaatg atgacaccaa gaaagatatg ctcggcaagc tgctgagcac gggctgg	900
cagaacttcc ctaacccat ccatatcaag ctatagagg gcaagttcca agacatgtg	960
aatatttcac agcatcagtg cgtcaagaag cagtgccccg aaaattctgg atgcttccgg	1020
cacctggatg agcgagaaga gtcagaatgc ctgcttaact ataaacagga gggcgacaaa	1080

-continued

tgtgtggaga acccaaatcc gacgtgcaac gagaacaacg gtggctgcga tgccgacgcg	1140
acttgtacag aggaagactc gggagttct cgaaaaaaa tcacgtgcga gtgcacccaa	1200
cccgacagtt atcctctgtt cgatggata ttctgctcct ccagcaacgt tactacttcc	1260
ggcactaccc gtcttctatc tggtcacacg tggcacgt tgacaggtt gcttggacg	1320
ctagtaacca tgggcttgct gacttaa	1347

<210> SEQ ID NO 8
<211> LENGTH: 1254
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: This is artificially synthesized MSP1 d-42* 3D7 sequence ORF

<400> SEQUENCE: 8

atgaccgtcg cgccggccgag cgtgcccccg gcgctgcccc tcctcgaaaa gctgccccgg	60
ctgctgctgc tggctgttt gtgcctgcgg gccgtgtgg gatccgttgtt caccggggaa	120
gcgattatgt tcactatggaa caaatatcc tcacgtgtttt agaaccggatc tgacgtgtatc	180
tacccatcaac cactagccgg agtttacaga agtctcaaga agcagatcga aaagaacatc	240
ttcaccttta atctaaacctt aaacgacatc ttgaattccc ggctgaaaaa gcggaaatac	300
ttccctcgacg tactggagtc ggatttgatc cagttttagt acatctccag caacgaatac	360
attatcgagg actcggttcaa actgtttaaac tccgagcaga agaacadcc gctgaagtcc	420
tacaaatata tcaaagagtc agtcgagaac gatattaaat tcgccccaga aggctatacg	480
tactacgaaa aggtccctcgcaaaatacaag gacgatctgg agtctatcaa aaaggtcatc	540
aaagaagaga aagagaaatt tcccgatctt ccccttacaa cgccggccctc tccagccaag	600
actgatgaac agaaaaaaa gtctaagttc ctccctttcc tcactaatat cgagacttc	660
tacaataacc tagtgaacaa gattgacgac tacctgatca accttaaagc caagataaac	720
gactgcaatg tcgagaagga tgaggctcat gttaagatca ccaactgtc cgatctgaaa	780
gccatcgacg acaagatcga cttatccaa accccatcgtt atttcgaggg tatcaaaaag	840
ctgatcaatg atgacaccaa gaaagatgt ctccggcaagc tgctgagcac gggctgttg	900
cagaacttcc ctaacaccat catatcaaag ctcatagagg gcaagttcca agacatgtg	960
aatatccac agcatcgtt cgtcaagaag cagtgccccg aaaattctgg atgcttccgg	1020
cacctggatc agcgagaaga gtgcaggatgc ctgcttact ataaacagga gggcgacaaa	1080
tgtgtggaga acccaaatcc gacgtgcaac gagaacaacg gtggctgcga tgccgacgcg	1140
acttgtacag aggaagactc gggagttct cgaaaaaaa tcacgtgcga gtgcacccaa	1200
cccgacagtt atcctctgtt cgatggata ttctgctcct ccagcaacgt tttag	1254

<210> SEQ ID NO 9
<211> LENGTH: 5013
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: This is artificially synthesized MSP1 d-190 FCB1 sequence ORF

<400> SEQUENCE: 9

atgaccgtcg cgccggccgag cgtgcccccg gcgctgcccc tcctcgaaaa gctgccccgg	60
ctgctgctgc tggctgttt gtgcctgcgg gccgtgtgg gatccgtgac ccacgaatcc	120

-continued

-continued

ctgctgttca atatccagaa caacattccc gttatgtatt ctatgttcga tagcctcaac	2580
aattctctct ctcactgtt catggagata tatgagaagg agatggctg caacctgtat	2640
aaaactcaaag acaacgacaa gattaagaac cttctggagg aagctaagaa ggtctccacc	2700
tctgttaaaa ctctctcttc cagctccatg caaccactgt ctctcacacc tcaagacaag	2760
cccgaaagtga gcgcctaaccga cgacacccctc cactcgacca accttaataa ctcactgaaa	2820
ctgtttgaga acatcctgtc tctcggcaag aataagaaca tctaccaaga acttattgga	2880
cagaaatcgt ccgagaacctt ctacgagaag atactgaaag acagcgacac attctataac	2940
gagagttca ctaacttcgt gaaatctaaa gccgatgata tcaactctct taacgatgaa	3000
tctaaacgta agaagctgga agaggacatc aataagctga agaagacact gcaactgagc	3060
ttcgacccgt acaacaagta ccaaactgaaa ctggagagac tcttcgacaa gaagaagaca	3120
gtcggcaagt ataagatgca gatcaagaag ttgactctgc tcaaggagca gcttgaagc	3180
aaactcaact cactgaacaa tccgaaacac gtactgcaga acttctcagt gttctcaac	3240
aagaagaagg aagccgagat cgccgagaca gagaacactc tggagaacac caagatttt	3300
ctcaaacact acaaaggccct cgtcaagtat tataatggcc agtcttcctc tctgaagact	3360
ctctccgagg agagcatcca gaccgaggat aactacgcca goctcgagaa cttcaaggc	3420
ctgtctaaagc tcgaaggcaa gctgaaggac aacctgaacc tggagaagaa gaagctcagc	3480
tacctctcta gcggactgca tcacctgatc gccgagctca aggaagtcat taagaacaag	3540
aactacaccc gcaatagccc aagegagaat aatacagacg tgaataacgc actggaatct	3600
tataagaagt tcctgcctga aggaacagat gtgcgcactg tgggtctga atctggctcc	3660
gacacactgg agcagtctca acctaagaag cctgcacatc ctcatgtcgg agccgagtcc	3720
aatacaatta ccacatctca gaacgtcgac gatgaggtcg atgacgtcat cattgtgcct	3780
atcttcggcg agagegagga ggactacgt gacctcgcc aggtggtcac cggagaggct	3840
gtcactcctt ccgtgattga taacattctg tccaaaatcg agaacgataa cgaagtgtc	3900
tatctgaaac ctctggcagg cgtctatagg tctctcaaga aacagctgga gaataacgt	3960
atgacccatca atgtcaacgt gaaggacatt ctgaacagcc gctttaataa gagagaaaaat	4020
ttcaagaacg tcttgagag cgacttgatt ccctataaag acctgaccc tcctaaactat	4080
gttgtcaagg acccatacaa gttcctcaat aaagagaaga gggataaatt tctgtctagc	4140
tacaactata tcaaggactc catgcacacc gatatcaatt tcgctaatga tggctgggg	4200
tattacaaga tcctgagcga aaaatacaag tctgacccctt actctattaa aaagtatatc	4260
aacgataagc aaggcgagaa tgaaaaatat ctgcccttcc tgaataacat cgaaaccctg	4320
tacaagacag tgaacgacaa aatcgaccc tcgtaattc acctggagcc caaggctcc	4380
aactatactt acgagaagag caatgtggaa gttaaaatca aggagctgaa ctacctcaaa	4440
acaatccaag acaagctggc agatttcaag aaaataaca atttcgtcgg aattgcagac	4500
ctgtctaccg attataacca caacaatctc ctgaccaagt ttctgtccac tggcatggtg	4560
ttcgaaaacc tcgccaaaac agtgcgtgac aatctgtcg acggcaaccc gctggccatg	4620
ctgaacatct cccagcacca atgcgtgaag aaacagtgcc cccagaatag cggctgttc	4680
aggcatctgg acgagcgcga agagtgcag tgcgtctgtc actacaaaca agaaggagat	4740
aagtgcgtgg agaaccctaaa ccctacctgc aatgaaaaca atggcggtg tgacgcccgt	4800
gtctaaatgca ccgaggaaga cagcggctc aacggaaaga aaatcacatg cgagtgtact	4860

-continued

aagcccgact cctatccact cttegacggg atcttctgct ccagctctag caacgttact	4920
acttcggca ctacccgtct tctatctggt cacacgttgt tcacgttgac aggtttgctt	4980
gggacgctag taaccatggg cttgctgact taa	5013

<210> SEQ ID NO 10
<211> LENGTH: 1119
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: This is artificially synthesized CS sequence
ORF

<400> SEQUENCE: 10

atgatgagga aactggccat cctgagcgtg agcagcttcc tgttcgtgga ggccctgttt	60
caggagtacc agtgctacgg cagcagcgc aacacccggg tgctgaacga gctgaactac	120
gacaacgccc gcaccaacct gtacaacgag ctggagatga actactacgg caagcaggag	180
aactggtaca gcctgaagaa gaacagccgg tctctgggccc agaacgcacga cggcaacaac	240
aacaacggcg acaacggccg ggagggcaag gacgaggaca agcgggacgg caacaacgag	300
gacaacgaga agctgcccggaa gcccaagcac aagaaactta agcagccccgc cgacggcaac	360
cccgacccca acgccaaccc caacgtggac cccaaacgcca atcctaattgt cgaccccaat	420
gccaatccga acgttgatcc caatgcgaat cctaacgcta accccaatgc caacccaaat	480
gccaatccaa atgcaaatcc caacgccaat ccaaacgcaa accctaattgc taatccaaac	540
gctaattctta atgccaatcc caatgctaac ccaaacgtcg atcctaacgc aaatccgaac	600
gctaaccctta acgccaatcc caacgctaac cccaaacgcaa accctaacgc caatccgaat	660
gccaacccaa acgccaaccc gaacgctaatt ccgaatgcta accccgaatgc taatccctaa	720
gcaaacccaa acgccaaccc caatgcaaac ccaaattgcca atcccaacgc caatcctaatt	780
gccaacaaga acaatcaggg caacggccag ggccacaaca tgcccaacga ccccaacccgg	840
aacgtggacg agaacgcca cggcaacacgc gccgtgaaga acaacaacaa cgaggagccc	900
agcgacaaggc acatcaagga gtacctgaac aagatccaga acagcctgag caccgagtg	960
agccctgca gcgtgacctg cggcaacggc attcaggatgc ggtcaagcc cggcagcgcc	1020
aacaagccca aggacgagct ggactacgcc aatgacatcg agaagaagat ctgcaagatg	1080
gagaagtgcgca gcagegtgtt caacgtgggta aactcctgta	1119

<210> SEQ ID NO 11
<211> LENGTH: 1734
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: This is artificially synthesized DiCo 1
complete sequence ORF

<400> SEQUENCE: 11

gttaccgtca cgcgtcaccg gtgtcatcat gaccgtggcc aggccctctg tgccctgcgc	60
cctggccctgt ctggggcagc tgccccggct gctgctctgt gtgctgttgt gcctgcccgc	120
cgtgtggggta tccgtgatcg agatcgtggta gcggagcaac tacatggca accccctggac	180
cgagtagatcg gccaagtacg acatcgagga agtgcacggc agcggcatcc ggggtggac	240
gggcgaggac gcccggatgtgg cccggacccca gtacaggatgc cccageggca agtgecccg	300
gttcggcaag ggcacatcatca tcgagaacag ccagaccacc ttccctgaccc ccgtggccac	360
cgagaaccag gacctgaagg acggcggtt cgccttcccc cccaccaacgc ccctgtatgag	420

-continued

ccccatgacc	ctggaccaga	tgccggactt	ctacaaggac	aacgagtacg	tgaagaacct	480
ggacgagctg	accctgtgca	gccggcacgc	cggcaacatg	aaccccgaca	acgacaagaa	540
cagcaactac	aagtaccccg	ccgtgtacga	cgacaaggat	aagaagtgcc	acatcctgta	600
tatcgccgccc	caggaaaaca	acggcccccag	gtactgcaac	aaggacgaga	gcaagcggaa	660
cagcatgttc	tgcttcagac	ccgccaagga	caagagcttc	cagaactacg	tgtacctgag	720
caagaacgtg	gtggacaact	ggggagaaagt	gtgcccccg	aagaatctgg	aaaacgc当地	780
gttcggcctg	tgggtggacg	gcaactgcga	ggacatcccc	cacgtgaacg	agttcagcgc	840
caacgacctg	ttcgagtgca	acaagctgtt	gttcgagctg	ttcgccagcg	accagccaa	900
geagtagcgg	cagcacctga	ccgactacga	gaagatcaa	gagggttca	agaacaagaa	960
cgccgacatg	atcaagagcg	ccttctgccc	aactggcgcc	ttcaaggccg	acagatacaa	1020
gagccacggc	aagggttaca	actggggcaa	ctacaacaga	aagaccaga	agtgc当地	1080
cttcaacgtg	aagccaccc	gcctgtatcaa	cgacaagtcc	tatatgc当地	ccaccgc当地	1140
gagccacccc	atcgaggtgg	agcacaactt	cccttgc当地	ctgtacaagg	atgagatcaa	1200
gaaagagatc	gagcgggaga	gcaagaggat	caagctgaac	gacaacgc当地	acgaggggcaa	1260
caagaagatc	attgc当地	ggatcttcat	cagcgacat	aaggacagcc	tgaagtgc当地	1320
ctgc当地	gagatcgtgt	cccagagcac	ctgcaat	ttcggtgca	aatgc当地	1380
gaagcgggccc	gaagtgtacca	gcaacaacga	ggtgtgtgt	aaagaggaat	ataaggacga	1440
gtacgc当地	atccccgagc	acaagccac	ctacgacaag	atgaagatca	tcattgc当地	1500
ctctgc当地	gtggccgtgc	tggccaccat	cctgtatgg	tacctgtaca	agcggaaagg	1560
caacgc当地	agatcgtata	agatggacca	gcctcagcac	tacggcaaga	gcaccagcc	1620
gaacgc当地	atgctggacc	ccgaggccag	cttctggggc	gaggaaaaga	gagctagcca	1680
caccacccccc	gtgctgtatgg	aaaagcccta	ctactgtatg	gccc当地	gctc	1734

<210> SEQ ID NO 12
<211> LENGTH: 1506
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: This is artificially synthesized DiCo 1 ecto sequence ORF

<400> SEQUENCE: 12						
ggtaccgtca	cgcgtaaccg	gtgtcatcat	gaccgtggcc	aggccctctg	tgcctgc当地	60
cctgccccctg	ctgggggagc	tgccccggct	gtgtctctg	gtgtgtgt	gcctgc当地	120
cgtgtgggg	tccgtgtatcg	agatcgtgg	gccccggcaac	tacatggca	accctggac	180
cgagtagatg	gccaagtacg	acatcgagga	agtgc当地	agcggcatcc	gggtggac	240
gggc当地	ggcggaggat	ccggcaccca	gtacaggctg	cccagggca	agtgc当地	300
gttc当地	ggcatcatca	tcgagaacag	ccagaccacc	ttcctgaccc	ccgtggccac	360
cgagaaccag	gacctgtatgg	acggccggtt	cgccctcccc	cccacca	ccctgtatg	420
ccccatgacc	ctggaccaga	tgccggactt	ctacaaggac	aacgagtacg	tgaagaacct	480
ggacgagctg	accctgtgca	gccggcacgc	cggcaacatg	aaccccgaca	acgacaagaa	540
cagcaactac	aagtaccccg	ccgtgtacga	cgacaaggat	aagaagtgcc	acatcctgta	600
tatcgccgccc	caggaaaaca	acggcccccag	gtactgcaac	aaggacgaga	gcaagcggaa	660
cagcatgttc	tgcttcagac	ccgccaagga	caagagcttc	cagaactacg	tgtacctgag	720

-continued

caagaacgtg gtggacaact gggagaaagt gtgcggccgg aagaatctgg aaaacgc当地	780
gttcggcttg tgggtggacg gcaactgc当地 ggacatcccc cacgtgaacg agttcagc当地	840
caacgacctg ttcgagtgc当地 acaagctggt gttcgagctg tccgccagcg accagcc当地	900
gcagtagcgg cagcacctg local ccgactacga gaagatcaaagg gagggttca agaacaagaa	960
cggccgacatg atcaagagcg ccttctgccc aactggcgcc ttcaaggccg acagatacaa	1020
gagccacggc aagggttaca actggggca local ctacaacaga aagaccaga agtgc当地	1080
cttcaacgtg aagccaccc gcctgatcaa cgacaagtcc tatategc当地 ccaccgc当地	1140
gagccacccc atcgagggtgg agcacaactt cccttgc当地 ctgtacaagg atgagatcaa	1200
gaaagagatc gagcgggaga gcaagaggat caagctgaac gacaacgc当地 acggaggcc当地	1260
caagaagatc attgc当地 ggttcat cagcgacgat aaggacagcc tgaagtgc当地	1320
ctgc当地 gagatcgtgt cccagagcac ctgcaatttcc ttctgtgc当地 aatgc当地	1380
gaagcgggccc gaagtgc当地 gcaacaacga ggtgggtggt aaagaggaat ataaggacga	1440
gtacgc当地 atcccgagc acaagccac ctacgacaag atgtgtatgat gagc当地	1500
gagctc	1506

We claim:

1. A combined measles-malaria immunogenic composition comprising a recombinant measles vaccine virus which expresses MSP1 malaria antigens and measles antigens capable of eliciting immune response against measles and malaria wherein the nucleotide sequence of MSP1 malaria antigen is selected from SEQ ID NO:7 and SEQ ID NO:8.

2. The combined measles-malaria immunogenic composition as claimed in claim 1 wherein the recombinant measles vaccine virus expresses single or different malaria antigens.

3. The combined measles-malaria immunogenic composition as claimed in claim 1 wherein the recombinant measles vaccine virus expresses MSP1 malaria antigen in both anchored and secreted forms.

4. The combined measles-malaria immunogenic composition as claimed in claim 1 wherein the recombinant measles vaccine virus expresses MSP1 malaria antigen in both anchored and secreted forms 3D7 strain and MAD 20.

5. The combined measles-malaria immunogenic composition as claimed in claim 1 wherein the recombinant measles vaccine virus expresses MSP1 malaria antigen in both anchored and secreted forms FCB1 strain.

6. The combined measles-malaria immunogenic composition as claimed in claim 1 wherein the recombinant measles vaccine virus expresses at least one of Diversity Covering (DiCo) AMA1 malaria antigen, DiCo-1 of AMA1 malaria antigen, DiCo-2 of AMA1 malaria antigen and DiCo-3 of AMA1 malaria antigen.

7. The combined measles-malaria immunogenic composition as claimed in claim 1 wherein the recombinant measles vaccine virus expresses DiCo-1, DiCo-2 and DiCo-3 of AMA1 malaria antigen.

8. The combined measles-malaria immunogenic composition as claimed in claim 6 wherein the recombinant measles vaccine virus expresses Diversity Covering (DiCo) of AMA1 malaria antigen in trans membrane and secreted forms.

9. The combined measles-malaria immunogenic composition as claimed in claim 1 wherein the recombinant measles vaccine virus expresses CS malaria antigen.

10. The combined measles-malaria immunogenic composition as claimed in claim 1 wherein the malaria antigen is cloned between P and M or H and L protein of recombinant measles vaccine virus.

11. A measles vaccine virus vector comprising the nucleotide sequence of an antigen of malaria wherein the nucleotide sequence is selected from SEQ ID NO:3 to SEQ ID NO:12.

12. A measles vaccine virus vector comprising the nucleotide sequence of an antigen of malaria, wherein the measles vaccine virus vector further comprises the nucleotide sequence selected from SEQ ID NO:1 and SEQ ID NO:2.

13. The vector as claimed in claim 11 wherein the nucleotide sequence encodes malaria antigens selected from d83-30-38 and d42 and d190 fragments of MSP1 or Diversity Covering (DiCo) AMA1 or CS protein.

14. A host comprising the vector of claim 11.

15. The host as claimed in claim 14 is selected from *E. coli* or mammalian cell line.

16. The combined measles-malaria immunogenic composition as claimed in claim 1 wherein the recombinant measles virus originates from a vaccine strain derived from Edmiston Zagreb.

17. A combined measles-malaria immunogenic composition comprising a recombinant measles vaccine virus which expresses MSP1 malaria antigens and measles antigens capable of eliciting immune response against measles and malaria, wherein recombinant measles vaccine virus comprises the following sequences:

MSP-1 d-190-3D7 SEQ ID NO:3
MSP-1 d-190*-3D7 SEQ ID NO:4
MSP-1 d-83-30-38-3D7 SEQ ID NO:5
MSP-1 d-83-30-38*-3D7 SEQ ID NO:6
MSP-1 d-42-3D7 SEQ ID NO:7
MSP-1 d-42*-3D7 SEQ ID NO:8
MSP-1 d-190-FCB1 SEQ ID NO:9
CS SEQ ID NO:10
DiCo1-complete SEQ ID NO:11
DiCo1-ecto and SEQ ID NO:12.

91

18. The combined measles-malaria immunogenic composition as claimed in claim **1** wherein recombinant measles vaccine virus further encoding a protein with adjuvantic properties.

19. The immunogenic composition as claimed in claim **1** further comprising an interleukin.

20. The immunogenic composition as claimed in claim **1** which comprises one of the described recombinant measles malaria viruses or a mixture of two to several such viruses.

21. The immunogenic composition as claimed in claim **1** wherein the described recombinant measles malaria viruses or a mixture of two to several such viruses devoid of defective interfering particles (DIs).

22. The immunogenic composition as claimed in claim **1** wherein the adventitiously arisen DI particles have been eliminated by plaque purification, by end point dilution or differential centrifugation.

23. The immunogenic composition as claimed in claim **1** being a component of a combined vaccine where the other

92

components are rubella, mumps, varicella or another live attenuated vaccine virus, naturally attenuated or recombinant, alone or in combination.

24. The immunogenic composition as claimed in claim **1** for parenteral administration comprising a suitable stabilizer.

25. The immunogenic composition as claimed in claim **1** comprising a suitable stabilizer, adjuvant or a combination thereof wherein the stabilizer and adjuvant are such that the vaccine can be administered parenterally, intranasally, by inhalation, orally, transdermally or in a suppository.

26. A composition comprising the combined measles-malaria immunogenic composition as claimed in claim **1** comprising stabilizer and/or adjuvant.

27. The immunogenic composition as claimed in claim **19** wherein the interleukin is interleukin 2.

28. The immunogenic composition according to claim **24** wherein the stabilizer is gelatin, human serum, albumin, sorbital or a combination thereof.

* * * * *